The Development of Systematic Controllability Assessment for Process Control Designs

This thesis is presented for the degree of
Doctor Philosophy in Engineering

by

Estiyanti Ekawati
B. Eng, M. Eng

School of Engineering Science
Murdoch University
December 2003

Declaration

I declare that this thesis is my own account of my research and contains as its main content work, which has not previously been submitted for a degree at any tertiary educational institution.

Estiyanti Ekawati

Abstract

Chemical process industries are constantly challenged to operate profitably and efficiently, despite the presence of significant uncertainties and disturbances on the operational conditions, and various operational limitations. The capability to meet the challenge relies on the quality of process control design, which should integrate the dynamic controllability characteristics in addition to the traditional economic considerations.

The focus of this thesis is the development of a systematic controllability assessment framework for process control design. The framework addresses the controllability aspects in process and controller structures, as well as in time-domain dynamic performances. The aim is to provide clearer relationships between process profitability, controllability, and operational switching strategies in response to variations in the operating conditions.

The skeleton of the framework is a mathematical optimisation algorithm. This algorithm considers the structural, operational and economic problems arising in process control design as a progressive, dynamic, and uncertain semi-infinite mixed integer nonlinear programming problem. The algorithm is an iterative, two-level optimisation, which determines the optimum process design and the associated controllability index within an optimisation window. The window progresses along a

time horizon, ensuring optimal process design within the window while accommodating the design switching during the course of load variations in a larger time horizon.

The controllability index quantifies the design capability to satisfy a given economic objective. Unique to other existing approaches, the process controllability index is computed based on the multi-dimensional geometric representation of the disturbances and uncertainties, measured process dynamics, and feasible operating spaces. These representations account for variable interactions existing in a multivariable process operation, in contrast to separate quantification in traditional single variable assessments.

The geometric computation of the index requires the analysis and elimination of redundant measurement variables, which occur in different combinations at different process and controller structures. The redundancy is detected and eliminated based on statistical collinearity among the process data, allowing the assessment to focus on the retained functional variables and the associated critical disturbances and uncertainties.

The redundancy analysis is tailored with a dynamic mixed integer nonlinear programming (MINLP) solver, which is dedicated to select the optimum process and controller structure within the design. The solver is developed based on the branch and bound strategy over the design tree, which consists of alternative nonlinear programming (NLP) sub-problems. In addition to the redundancy analysis, the solver is equipped with a compact MINLP formulation, an alternating depth-first and breadth-first search strategy, and the rapid pruning of inferior NLP sub-problems. The tailored strategy ensures fast and efficient convergence of convex problems, as well as superior optimum of non-convex counterparts.

Finally, the framework is performed within a time window, which progresses along the time horizon. This strategy provides realistic responses to major variations along

greater length of time, by switching between optimum operational modes, while maintaining the optimum process controllability.

The performance of the framework is illustrated through several case studies. Each case demonstrates the novelty of addressing various computational features in a concise algorithm. These include the industrial case, which involves the systematic controllability assessment of an industrial five-effect liquor-burning evaporator within an Alumina refinery, which highlights the contribution of this framework in bridging the process design methodologies with the industrial implementation.

The thesis consists of eight chapters, presenting the systematic development of the framework. The numerical implementations have been organised in a MATLAB Toolbox, accompanied with the relevant case studies.

Acknowledgements

Four years of carrying out this research at the School of Engineering Science, Murdoch University, Australia, with financial support from the Australian Development Scholarship, has been a fantastic experience. This is a marvelous place to study, with excellent facilities and wonderful people. It is a great pleasure to acknowledge this institution, as well as the following people who have made this endeavor possible.

The sincerest gratitude goes to Dr. Parisa Arabzadeh Bahri, who has been the most inspiring and generous supervisor. Her insights into process controllability, as well as her research management and publication capabilities, have established the direction of this research and made it an enjoyable journey. Her support and encouragement have been exceptional both in professional and personal levels. I am honored and privileged to be one of her PhD students.

I would also like to express the great appreciation to Mr. Graham LePage, Dr. Ali Nooraii, and the Editorial Board of Alcoa Australia Limited, for their guidance and professional assistance in the industrial case study of this research.

My appreciation extends to the fellow and former postgraduate students and staff of the School of Engineering Science, as well as the staff of the Rockingham Regional Campus Community Library and Murdoch International for their friendship and

excellent support. In particular, I would like to thank Daniel McGill, Dr. S. Srinivas Shastri, Mahsa Ghaeli and Janet Blinston for their unrelenting encouragement during the hardest part of thesis writing.

Lastly, I dedicate this work to the most important people in my life; my husband Eko Mursito Budi, my children Karismanto Rahmadika and Krisna Diastama, and our parents. Their faith on me, infinite love and patience had inspired me at every steps of this research.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Chemical process industries, such as mineral plants in Australia, are subject to significant variation in their operational conditions. The variations may occur due to external disturbances, such as fluctuations of quality and amount of feed material, fuel, as well as utility streams; or process uncertainties, such as unmeasurable parameters, or unexplainable dynamics within the process. Despite the loads, the market competition demands for a minimum variation in product quality. In addition, the industries are also under pressure to satisfy the economic objective and environmental legislation. These lead to higher material and energy utilisation, which are implemented in utility sharing, stream recycling, and numerous interconnections. The resulting processes are commonly complex and highly interactive.

In response to those challenges, the industries have developed a greater appreciation of process control design. It is accepted that a well-designed process control yields optimum performances, specifically profitability and dynamic operability. This trend is supported by intensive development of process control methods, which produce various strategies, from Proportional-Integral (PI) control, to Model Predictive Control (MPC) schemes.

Facing different techniques, it seems natural to inquire which is the best plant design and control strategy for a given economic objective. This problem calls for the rigorous analysis of process controllability, which is the ability of the plant to achieve a desired dynamic performance within various limitations on process operation, despite the external disturbances and uncertainties in design parameters, using the available inputs and manipulated variables (Skogestad and Wolff, 1992).

Traditionally, process control design has been a sequential procedure. Firstly, the structure and parameters of the process were designed to satisfy economic steady state objectives subject to system constraints. Subsequently, the controller structure and parameters were designed to achieve a desired dynamic performance. This approach could trap the designer to select the cheapest design that is difficult to control and therefore requires expensive over-design; leaving out a slightly less economic design that is easier to control and in overall more profitable (Weitz and Lewin, 1996).

This condition highlights the interaction between the process design and the selected controller strategy. Even in early days of control studies, Ziegler and Nichols (1943) observed that a poor controller could perform acceptably on an easily controlled process. In contrast, the finest controller might never deliver the desired dynamic performance when applied to a miserably designed process. It is quite interesting that over six decades later this is still a common problem in process design! Therefore, one could never over-emphasize the importance of integrating controllability assessment with process design at the earliest possible stage (Bahri *et al.*, 1996).

For that purpose, numerous controllability assessments have been proposed in the open literature. However, these researches offered partial solutions, which can be grouped as follows:

1. The analyses based on linear process model and quantified in frequency domain (Hovd and Skogestad, 1994; Morari and Skogestad, 1985; Wolff *et al.*, 1992; Zhao and Skogestad, 1997).
2. The analyses of process structure only (Georgiou and Floudas, 1989; Hopkins *et al.*, 1998), and the analyses of the dynamic performance of fixed process and

controller structures only (for instance, Chenery, 1997), which do not provide clear relationship between process economy, controllability and design.

3. The analyses subject to simple disturbance profiles, such as scalar, step or sinusoidal functions (Bahri, 1996; Narraway and Perkins, 1994; Schweiger and Floudas, 1998).

Linear controllability analyses are computationally tractable, and commonly provide elegant analytical expressions in frequency domain. However, these techniques require linearised process models around the steady state operating points. Therefore, they are only reliable within narrow operating regions of real, inherently nonlinear processes. Accordingly, nonlinear dynamic performances are not represented accurately in the frequency domain. This also applies for process constraints and disturbance profiles. In contrast, while there is no guarantee of general and elegant solutions, the time domain accommodates these controllability aspects, and provides the ease of interpreting the results. Therefore, when accompanied with a strong numerical solution, the time domain is considered as a suitable field to analyse process controllability (Bequette, 1991; Chenery, 1997).

The analyses focusing on process structure only, in general provide short-listed design alternatives based on the interrelationship between process variables. These require further economic assessments to determine the best structure; hence, these are not comprehensive solutions by themselves. On the other hand, the analyses of a fixed process and controller structure may yield economically optimum solutions over a set of parameter values. However, the process and controller structure under study may not be the best design. Therefore, controllability analysis should screen the feasible process structures, and determine the rank of the control performances simultaneously with economic assessment (Bahri *et al.*, 1996).

Considering the above limitations, it is argued that mathematical programming, or the optimisation method, is the most suitable approach to accommodate simultaneous assessments of process economy and controllability of general nonlinear systems. The process descriptions, performance specifications, and limitations on control

performances in time domain, as well as the process and controller structure selection, are all accommodated within the problem formulation (Chenery and Walsh, 1998). It is acknowledged that the cost of numerical computation is considerably higher than the linear analysis. However, this load can be shared with the recent advances of the computer technology and the associated increase of computation speed.

In combination with the optimisation approach, the quantification of process controllability would establish a clearer relationship between the economic and dynamic performance objectives, and provide more sound advice for the management decision. One example of such a quantification is the required amount of back off from the process operational limits to ensure feasible dynamic operation (Narraway *et al.*, 1991). This concept notes that the optimum steady state operating point of a process, which is when the process is not affected by disturbances and uncertainties, would be located on the operational limits. The process in this situation can easily become infeasible when disturbed. Therefore, one should find another operating point, if any exists at all, that keeps the process feasible despite the presence of the disturbances and uncertainties, and as close as possible to the nominal steady state point.

Based on this concept, Bahri (1996) introduced the original Dynamic Operability Framework, an iterative dynamic optimisation algorithm to assess the flexibility and controllability of linear and nonlinear processes. Within this framework, the process dynamics are assessed against the disturbances and uncertainties, which are assumed as sets of step changes from given nominal values. The step magnitudes are assumed uniformly varying between the upper and lower bounds. This assessment compares the steady state operating point, which is typically located on the operational constraints; with the dynamic operating points, which are located further inside the feasible operational space. The back off movement is defined as the differences between the steady state and the dynamic profits (costs) at the respective operating points. Accordingly, it becomes the economic index of the process.

This index not only depends on the process model, but also on the magnitudes of disturbances and uncertainties, and the implemented controller. Therefore, it principally can be used to rank several alternative control strategies. However, the index only provides the difference between two nominal conditions. The only regulatory performance deduced from this information is the maximum offset or the overshoot of the worst-case dynamics. For instance, consider two types of regulatory dynamics. The first one has persistent offset. The other has aggressive control, with overshoot of the same magnitude, but quickly settles back to the original course. The back off index obviously cannot distinguish between these dynamic conditions. Therefore, the index requires extension to cover the overall regulatory dynamic profile.

One well-known alternative for quantifying the dynamic profile is the Integral Squared Error (ISE, Schweiger and Floudas, 1997), or its variant the Integral Absolute Error (IAE). However, ISE only represents one profile at a time. For a multivariable process with several measured output variables, it is not clear which variables should be assessed with ISE. This is an important decision, because different process structures may activate different dynamic profiles and different constraints. If the variable assessed with ISE happens to be inactive at a certain structure, for instance, because the corresponding equipment does not exist, then the ISE is either nil or undefined. The first case would easily mislead the controllability analysis. Furthermore, ISE only quantifies the dynamic profile, and does not guarantee the feasible operation.

These conditions show that a controllability index to guarantee feasible operation and quantify multivariable controllability is yet to be developed, and worthy of further investigation.

Finally, the reported controllability assessments are subject to simple disturbance profiles (Bahri, 1996; Narraway and Perkins, 1994; Schweiger and Floudas, 1998). The exception is found in Chenery (1997), where the controllability of linear, fixed structure process model is assessed subject to a general disturbance profile. The assessment subject to simple profile may fail to anticipate broader possible dynamics within the operating domain, or on the contrary, become too conservative. Therefore, the

controllability assessment involving nonlinear model and structure selection, subject to general disturbance profile requires further work.

It has been shown that a comprehensive methodology to assess the process controllability deserves an active investigation. The method is to be formulated and interpreted in time domain, addressing the controllability of both process structures as well as dynamic performances, providing a multivariable quantification of process dynamics, and anticipating broader type of variation in operating conditions. These aims are formally stated in the following section.

1.2 Objectives

This thesis focuses on the interaction between process design and control. Accordingly, the objective is to develop a systematic methodology to assess the controllability of process control design, in form of an optimisation framework.

This integration should deliver the following properties:

1. A comprehensive analysis of process controllability.
   The analysis evaluates the process and controller structures as well as the parameters to optimise a given economic objective, whilst maintaining feasible operating conditions subject to general disturbance profiles.

2. The realisation of the analysis in a systematic optimisation framework with the following features:
   a. The quantification of multivariable dynamic performances to represent process controllability.
   b. The solution of dynamic MINLP problem for a given economic objective.
   c. The accommodation of general disturbance profiles.

3. Implementation on several academic process models and a real industrial case. The industrial case study involves a five-effect liquor burning evaporator associated with the Bayer process in an Alumina refinery. The case study contributes to process industries in terms of the recommendation of optimum process control design, also the process structure that minimises the operational cost and its variability.

1.3 Scope

This thesis revisits the controllability assessment features in the original Dynamic Operability Framework (Bahri, 1996) and provides the required enhancement to cover the objectives stated in section 1.2.

The original framework is a two-level optimisation algorithm, addressing the feasibility and controllability assessment as semi-infinite dynamic Mixed Integer Nonlinear Programming (MINLP) problems. The processes under study are represented by a set of nonlinear Differential Algebraic Equations (DAE). The disturbances are assumed as a set of step functions with magnitudes uniformly varying within upper and lower values, from given nominal values. The second level of the algorithm automatically determines the critical disturbances, which transform the semi-infinite optimisation problem into a deterministic multi-period problem. The original framework is solved in GAMS software (Brooke *et al.*, 1992), where the DAE is solved using the orthogonal collocation on finite elements method (Vassiliadis *et al.*, 1994), and the MINLP problem with DICOPT (Viswanathan and Grossmann, 1990).

This research develops an alternative and enhanced methodology, maintaining the optimisation approach. Here, the semi-infinite dynamic MINLP is modified to incorporate a multivariable controllability index, a sequential DAE solution, a new MINLP solution, and the general disturbance profiles. These developments combine the knowledge of control theory, process design, mathematical programming, computational geometry, and statistics; supported by the vast and diverse literature

reviews within the thesis. Each development stage associates with one enhanced feature, involving concise formulations and discussion of practical realisations. These stages have also been reported in various conference and journal papers.

The numerical solutions are implemented using the technical software MATLAB ver 6.0 (Mathworks, 2000). Should non-MATLAB codes are required, appropriate interfaces would be developed to enable direct access from MATLAB environment. The computational performances at various stages of the development are validated and improved based on four academic models and one industrial case study.

Overall, the scope of this research is summarised in the Research Design and Documentation Diagram in Figure 1.1.

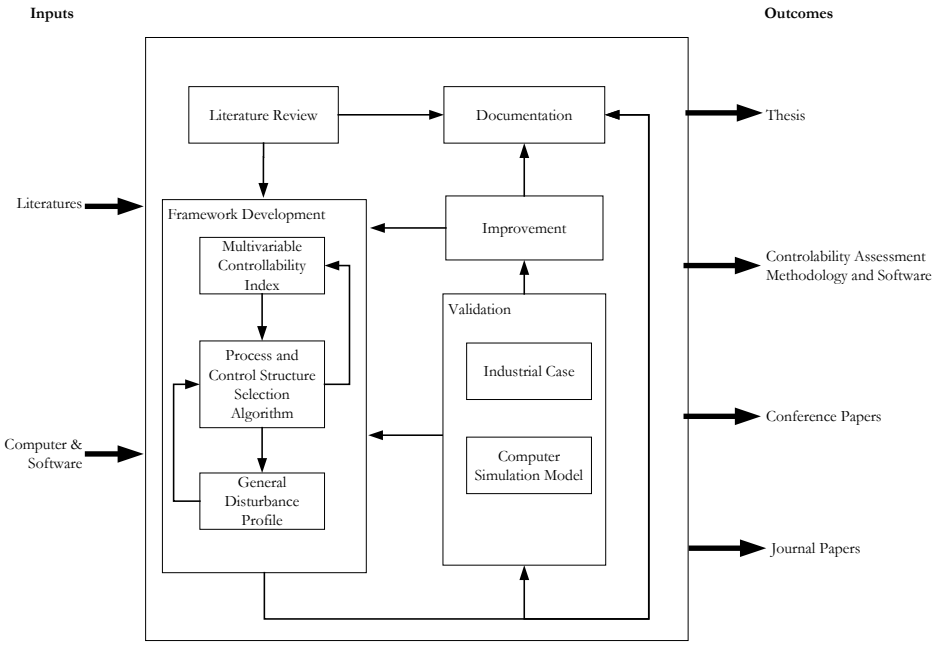


Figure 1.1 Research Design and Documentation Diagram

1.4 Thesis Structure

The chronology of the framework development strongly shapes the organisation of this thesis. So far, this chapter has presented the background, objective and scope of the research. It concludes with the description of thesis structure.

Chapter 2 presents an overview of the existing strategies on process operability, especially the flexibility and controllability issues. It is highlighted that the controllability considerations lead to process designs that are not only economically viable, but also controllable. These considerations can be accommodated systematically at the process design stage with the use of optimisation algorithms. Independent controllability measurements for linear and nonlinear models are reviewed. These are followed by the discussions of their integration into the process operability framework, especially those using the optimisation algorithm, including the original Dynamic Operability Framework (Bahri, 1996). The discussion then focuses on the potential benefits gained from accommodating a multivariable controllability index, dynamic MINLP solutions, and general characterisation of disturbances within the framework. The presentation of strategies to acquire these features finalises this chapter.

Chapter 3 presents the integration of the regulatory Output Controllability Index (r-OCI) within the Dynamic Operability Framework (DOF) for assessments of fixed-structure regulatory process control design. This index utilises the geometric representation of a feasible operating region, and the projection of disturbance space into output space to quantify the controllability. Within this approach, flexibility becomes the special case of the controllability problem. The integrated framework is applied on a nonlinear chemical process system with a fixed control structure.

Chapter 4 addresses the problem of geometric computation of the controllability index, when some of the measured variables are collinear. This collinearity relates to redundant dimensions in variable space, which disrupts the geometric computation, and causes incorrect detection of critical disturbance and uncertainty combinations within the framework. In this chapter, the redundant variables are detected and

eliminated based on the correlation Principal Component Analysis (PCA, Jolliffe, 1986) of steady state process data. This leads to the extraction of the minimum number of critical disturbances, which assists the correct quantification of controllability, and significantly improves the computational cost.

Chapter 5 extends the framework to accommodate the process and controller structure selection. Here the framework is formulated as a dynamic semi-infinite mixed integer nonlinear programming (MINLP) problem. The corresponding dynamic MINLP solution is developed based on the Branch and Bound (BB) method. The features include the accommodation of redundancy analysis and elimination procedure developed in Chapter 4, the compact MINLP superstructure formulation, the combined depth-search and breadth-search within the branching rules, and the convergence treatment of the NLP subproblems. The algorithm performance is demonstrated through four case studies, including the process synthesis of dynamic and uncertain nonlinear process in a superstructure.

Chapter 6 presents the controllability assessment subject to general disturbance profiles, focusing on the measurable disturbances. This is realised by sampling the disturbance profiles, and accordingly evaluating the associated piecewise process responses within the framework. To further reduce the conservatism in the design, the framework has been modified to progress sequentially in an optimisation window over the time horizon. In this sequential framework, the design in a sequence is based on the analysis of the sampled disturbances and the associated piecewise process responses within the preceding optimisation window. The performance of the framework is demonstrated through a case study involving a nonlinear process in a superstructure, which is affected by various disturbance signals.

Chapter 7 reports the application of the enhanced framework for systematic controllability assessment of an industrial five-effect liquor-burning evaporator associated with the Bayer process in an Alumina refinery. The process is characterised by strong interaction between dynamic variables, and an index-2 Differential Algebraic Equations (DAE) model. The process is considered within a superstructure, involving

process and controller structure selection. The controllability assessment covers both the worst-case and the general disturbances profiles over 24-hours operational period. The outcomes are the operational structure and the multi-loop control strategy implemented to achieve the optimum process economy.

Finally, Chapter 8 summarises and concludes the extent of this thesis to answer the objectives stated in this chapter. It reiterates the contribution of the research and recommending several potential further works in this area.

Each chapter begins with a brief abstract, outlining the content. Since the content of each chapter is quite specific, every chapter begins with a review of relevant literatures. A dedicated conclusions section is given after the presentation of the main subject and the case studies in each chapter. The chapters are also completed with nomenclatures and list of references.

The earlier versions of the chapters have been published in various conference and journal papers. The list of these publications has been provided at the beginning of the thesis, and each would be mentioned at the beginning of the relevant chapters. Within this thesis, these publications have been rewritten and reorganised. The general literature review has been moved to Chapter 2, such that the reviews in the other chapters can be more specific.

1.5 Conclusion

This thesis expresses the interest on the interaction between process design and process control, by developing the controllability assessment framework for process control design. The assessment would cover the general nonlinear multivariable regulatory dynamics, quantify the associated controllability and accommodate the general disturbance profiles. The development of the framework is described and analysed step-by-step in the next six chapters, and the thesis is concluded in Chapter 8.

1.6 Nomenclature

Acronyms		
DAE	:	Differential Algebraic Equations
DOF	:	Dynamic Operability Framework
IAE	:	Integral Absolute Error
ISE	:	Integral Squared Error
MINLP	:	Mixed Integer Nonlinear Programming
MPC	:	Model Predictive Control
NLP	:	Nonlinear Programming
OA	:	Outer Approximation
ODE	:	Ordinary Differential Equations
PCA	:	Principal Component Analysis
PI	:	Proportional Integral
r-OCI	:	Regulatory Output Controllability Index

1.7 References

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Bahri, P. A., Bandoni, A. and Romagnoli, J. (1996). "Operability Assessment in Chemical Plants", Computers and Chemical Engineering, 20, S787 - S792.

Bequette, B. W. (1991). "Nonlinear Control of Chemical Processes: A Review", Industrial and Engineering Chemistry Research, 30, 1391 - 1413.

Brooke, A. D., Kendrick and Meeraus, A. (1992). "GAMS 2.25 User's Guide".

Chenery, S. D. (1997). "Process Controllability Analysis Using Linear and Nonlinear Optimisation", Ph.D. Thesis, Centre for Process System Engineering and Department of Chemical Engineering, Imperial College of Science, Technology and Medicine, London.

Chenery, S. D. and Walsh, S. (1998). "Process Controllability Analysis Using Linear Programming", Journal of Process Control, 8, 165 - 174.

Georgiou, A. and Floudas, C. A. (1989). "Structural Analysis and Synthesis of Feasible Control Systems, Theory and Applications", Chemical Engineering Research and Design, 67, 600 - 618.

Hopkins, L., Lant, P. and Newell, R. B. (1998). "Output Structural Controllability: A Tool for Integrated Process Design and Control", Journal of Process Control, 8 (1), 57 - 68.

Hovd, M. and Skogestad, S. (1994). "Pairing Criteria for Decentralized Control of Unstable Plants", Industrial and Engineering Chemistry Research, 33, 2134 - 2139.

Jolliffe, I. T. (1986). "Principal Component Analysis", New York, Springer-Verlag.

Mathworks (2000). "Using MATLAB Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Morari, M. and Skogestad, S. (1985). "Effect of Model Uncertainty on Dynamic Resilience", I. Chem. E. Symposium Series.

Narraway, L. and Perkins, J. (1994). "Selection of Process Control Structure Based on Economics", Computers and Chemical Engineering, 18 (Supplementary), S511-S515.

Narraway, L. T., Perkins, J. D. and Barton, G. W. (1991). "Interaction between Process Design and Process Control: Economic Analysis of Process Dynamics", Journal of Process Control, 1, 243 - 250.

Schweiger, C. A. and Floudas, C. A. (1997). "Interaction of Design and Control: Optimization with Dynamic Models", Optimal Control: Theory, Algorithms and Applications, Hager, W. W. and Pardalos, P. M., Kluwer Academic Publishers B. V.

Schweiger, C. A. and Floudas, C. A. (1998). "Process Synthesis, Design and Control: A Mixed-Integer Optimal Control Framework", IFAC Conference on Dynamics and Control of Process Systems, Corfu, Greece, Elsevier.

Skogestad, S. and Wolff, E. A. (1992). "Controllability Measures for Disturbance Rejection", IFAC workshop on Interactions between Process Design and Process Control, London.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Viswanathan, J. and Grossmann, I. E. (1990). "A Combined Penalty Function and Outer Approximation Method for MINLP Optimisation", Computers and Chemical Engineering, 14, 769-782.

Weitz, O. and Lewin, D. R. (1996). "Dynamic Controllability and Resiliency Diagnosis Using Steady State Process Flowsheet Data", Computers and Chemical Engineering, 20 (4), 325 - 335.

Wolff, E. A., Skogestad, S., Hovd, M. and Mathisen, K. W. (1992). "A Procedure for Controllability Analysis", IFAC workshop on Interactions between Process Design and Process Control.

Zhao, Y. and Skogestad, S. (1997). "Comparison of Various Control Configuration for Continuous Bioreactors", Industrial and Engineering Chemistry Research, 36, 697-705.

Ziegler, J. G. and Nichols, N. B. (1943). "Process Lags in Automatic Control Circuits", Transactions of ASME, 65, 433.

Chapter 2

Operability Analysis in Process Control Design

2.1 Introduction

This chapter reviews the range of the existing operability analyses, specifically those leading to the incorporation of controllability assessments within process control design and synthesis. The objective is to present the background knowledge, starts from the description of process properties; to the operability properties, especially the flexibility and controllability; and concludes with the integrated economic and controllability assessments in nonlinear regulatory process design and synthesis.

The remainder of the chapter is organised as follows: Section 2.2 introduces and defines the variables and their relationships within a process control scheme. These definitions are used consistently throughout this thesis. Section 2.3 follows with the discussion of integrated attitude towards process design and control design. Section 2.4 describes the operability aspects of concern in process design, specifically the flexibility and controllability. The use of optimisation methods to integrate these aspects in process design is discussed in section 2.5. Sections 2.6 and 2.7 review the flexibility and controllability studies respectively. Section 2.8 discusses the integration of the

economic and controllability properties in process designs. The proposed development is discussed in section 2.9, and finally, the chapter is concluded in section 2.10.

2.2 Process Description

We may consider the process of filling a tank with liquid to be as simple as the change of liquid level, or as complicated as the mixing and turbulence patterns caused by the force of incoming stream. It shows that the scope of the process behavior is only limited by the point of view of the observer. In many levels of complexity, process behavior can be represented by a mathematical model, which is derived using the mass and energy balances combined with relevant physical and chemical properties. This research represents the model using a set of ordinary differential equations (ODE) or differential algebraic equations (DAE), either in implicit or explicit forms as follows:

$$\mathbf{h_i}(\mathbf{r},\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}_e,\boldsymbol{\theta}_h,\mathbf{p},t)=0 \qquad \mathbf{i}\in\mathbf{E} \tag{2.1}$$

$$\frac{d\mathbf{x}}{dt}=\dot{\mathbf{x}}=\mathbf{f_{i_1}}(\mathbf{r},\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta}_e,\boldsymbol{\theta}_h,\mathbf{p},t) \qquad \mathbf{i}_1\cup\mathbf{i}_2=\mathbf{i} \tag{2.2}$$

$$0=\mathbf{f_{i_2}}(\mathbf{r},\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta}_e,\boldsymbol{\theta}_h,\mathbf{p},t) \qquad \mathbf{i}\in\mathbf{E}$$

The expression (2.1) is the implicit form, where $\mathbf{h_i}$ represents the process model with $\mathbf{i}\in\mathbf{E}$ number of equalities. Alternatively, the explicit form (2.2) is also commonly used, with $\mathbf{h_i}$ transformed into $\mathbf{f_{i1}}$ and $\mathbf{f_{i2}}$. An ODE consists of $\mathbf{f_{i1}}$-type equations only, while DAE consists of both $\mathbf{f_{i1}}$ and $\mathbf{f_{i2}}$. In both expressions, **x** and **u,** are the vectors of state, and manipulated variables respectively. The vector of state derivative $\dot{\mathbf{x}}$ represents the process dynamics, which are observed through the measured outputs **w**. The dynamics are due to the changes over time **t** in the reference **r**, the external disturbance $\boldsymbol{\theta}_e$ and the process uncertainty $\boldsymbol{\theta}_h$.

The reference signals **r** dictate the corresponding output profiles. If the reference changes within a time horizon of interest, the capability to closely tracing the change is called the servo, or tracking performance. Meanwhile, a fixed reference is called a set point; and the capability to stay close to the point is called the regulatory performance.

In contrast, the dynamics due to $\boldsymbol{\theta}_e$ and $\boldsymbol{\theta}_h$ are unfavorable, since these may drive the process away from the intended profiles. Accordingly, a control strategy is required to compensate for these effects. This strategy drives the manipulated variable **u** based on the measurement of the output variable **w**, and may as well the require measurement of the disturbance $\boldsymbol{\theta}_e$.

The mathematical relationships between **r**, **u**, and $\boldsymbol{\theta}$ with **w** can be derived from (2.1) or (2.2), and accordingly referred to as the transfer functions. One can use these functions to evaluate the process dynamic performances and the corresponding control strategies.

The fixed variables during process dynamics are the constant process parameter **p**, the continuous design variable **z** and the discrete design variable **y**. The common examples of these variables are the equipment sizes and existences. Once decided in process design, they are fixed during process operations. The relationships between these variables are shown diagrammatically in Figure 2.1.

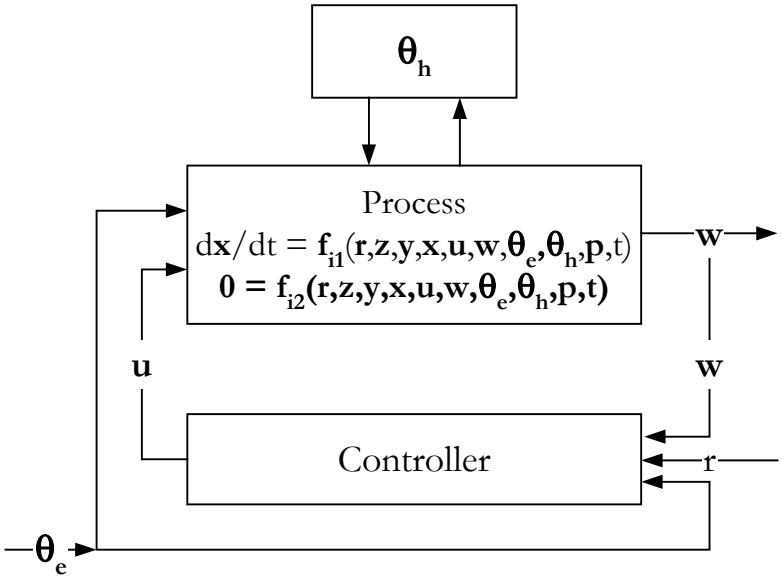


Figure 2.1 Basic Process Control Diagram

This research focuses on the regulatory processes. It is assumed that the set points are the function of the design variables **z** and **y**; hence **r** is omitted in the later descriptions. When there is no requirement to distinguish the unwanted $\boldsymbol{\theta}_e$ and $\boldsymbol{\theta}_h$, symbol $\boldsymbol{\theta}$ is used. The constant parameter **p** may also be omitted. Hence, the process dynamics may be given, either implicitly or explicitly, as follows:

$$\mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta},t)=0 \qquad \mathbf{i}\in\mathbf{E} \tag{2.3}$$

$$\frac{d\mathbf{x}}{dt}=\dot{\mathbf{x}}=\mathbf{f}_{\mathbf{i}_1}(\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta},t) \qquad \mathbf{i}_1\cup\mathbf{i}_2=\mathbf{i} \tag{2.4}$$

$$0=\mathbf{f}_{\mathbf{i}_2}(\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta},t) \qquad \mathbf{i}\in\mathbf{E}$$

Furthermore, the process dynamics may be constrained within the feasible operational spaces as follows:

$$\mathbf{g_j}(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta},t)\le\mathbf{0} \qquad \mathbf{j}\in\mathbf{I} \tag{2.5}$$

$$\mathbf{z}\in\mathbf{Z}=(\mathbf{z}\mid\mathbf{z}^{\mathbf{L}}\le\mathbf{z}\le\mathbf{z}^{\mathbf{U}})$$
$$\mathbf{y}\in\{0,1\}$$
$$\mathbf{x}\in\mathbf{X}=(\mathbf{x}\mid\mathbf{x}^{\mathbf{L}}\le\mathbf{x}\le\mathbf{x}^{\mathbf{U}})$$
$$\mathbf{u}\in\mathbf{U}=(\mathbf{u}\mid\mathbf{u}^{\mathbf{L}}\le\mathbf{u}\le\mathbf{u}^{\mathbf{U}})$$
$$\mathbf{w}\in\mathbf{W}=(\mathbf{w}\mid\mathbf{w}^{\mathbf{L}}\le\mathbf{w}\le\mathbf{w}^{\mathbf{U}})$$
$$\boldsymbol{\theta}\in\boldsymbol{\Theta}=(\boldsymbol{\theta}\mid\boldsymbol{\theta}^{\mathbf{L}}\le\boldsymbol{\theta}\le\boldsymbol{\theta}^{\mathbf{U}})$$
$$t=(t\mid t_o\le t\le t_f)$$

Here, the set of functions $\mathbf{g_j}$ with **I** number of inequalities defines the feasible operating region of the process dynamics. This set is associated with the performance requirements, equipment limitations and safety regulations. The variable ranges are defined in **X, U, W,** and **Θ,** with subscripts **L** and **U** as the lower and upper bounds, respectively.

There are also studies focused on steady state process conditions, where the dynamics have been diminished. In this case, the derivatives $\dot{\mathbf{x}}$ and time **t** are omitted from the model. The algebraic equalities are therefore simplified further by defining **u** and **x** in terms of **w**, **y**, **z** and $\boldsymbol{\theta}$. This gives the steady state model in (2.6).

$$\mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{w},\boldsymbol{\theta}) = 0 \qquad \mathbf{i} \in \mathbf{E} \tag{2.6}$$

The analyses of process dynamics and steady states are commonly performed against one or more objectives $\boldsymbol{\Phi}_q$, such as operational cost or dynamic performances. Based on the ODE/DAE description, these objectives can be regarded as the subset of **w**. However, since achieving the objectives is the goal of process design, they are formulated specifically as the functions of process variables. The expressions are given in (2.7) and (2.8) for dynamic and steady state processes, respectively.

$$\boldsymbol{\Phi}_q(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta},t) \qquad q = 1,2,\ldots \tag{2.7}$$

$$\boldsymbol{\Phi}_q(\mathbf{z},\mathbf{y},\mathbf{w},\boldsymbol{\theta}) \qquad q = 1,2,\ldots \tag{2.8}$$

The above process descriptions and definitions are the main vocabularies of the discussion about process control designs in the following sections.

2.3 Interaction between Process Design and Process Control

During the past three decades, the awareness of the importance of safety, profitability and efficiency of process systems has motivated significant number of studies about process control designs. It has been accepted that a well-designed process control system provides better performances, such as economic profit, safety and favorable dynamics. Considering the intensive development in process control theories, it is not surprising that one may face several different designs and controller strategies for one particular process. This leads to questions such as; 'How to assess these different designs?', 'How to ensure a particular choice is the optimal process control design?', or 'Which one should be handled first, the process, or the controller ?'

Traditionally, processes and controllers are designed sequentially. Firstly, the process configurations (structures) and parameters are designed to satisfy the economic objectives, such as maximum profits or minimum operational costs. The designs are

based on steady state models, and subjected to the operational constraints. Afterwards, the controllers are designed to reject the likely effects of external disturbances and process uncertainties, as well to achieve the desired dynamic performance. This approach carries a risk in that it may end up choosing the cheapest process design that was difficult to control. It may also miss out a slightly less economic but easier to control designs, the one that might be more profitable in a long run (Weitz and Lewin, 1996).

The above condition is the result of interactions between process design and process control. Marlin (1995) warns that the process control dynamic performances are significantly influenced by the operating conditions selected in the process design. This applies to chemical processes, due to their non-linear behavior. The dynamics may vary significantly at different operating conditions. Consequently, different control actions, even different control structures, may be required to maintain the desired performances. However, certain process configurations and operating conditions may be inherently uncontrollable. This fact is easily overlooked if the process is designed based on steady state economics only. Therefore, the scope of the design should be extended to include the proper assessment of the process dynamic performances. This assessment should cover both process and controller, at the respective structures and parameters, in the early stage of process design (Bahri *et al.*, 1996b).

2.4 Process Operability

How significantly does the consideration of process dynamics affect the quality of a process control design? This question can be addressed based on the operability properties. These include flexibility, controllability, optimality, stability, selection of measurements and manipulated variables (Wolff *et al.*, 1992a).

The flexibility, or feasibility, has been the most studied operability property in open literature of chemical process designs during the last three decades. It was defined as 'the ability to maintain the process variables within feasible operational region, despite

the presence of uncertainties' (Grossmann *et al.*, 1983). Flexibility is often sought after simultaneously with the economic objectives, hence raises the optimality issue. Therefore, flexibility studies are dominated by numerous optimisation strategies. Those studies aim at the determination of flexible operational spaces and flexibility measurements. Grossmann *et al.* (1983) are among the first to present the fundamental formulations of both aims within optimisation problems.

Until the mid-1980s, the flexibility studies were limited to steady state conditions, due to the constraints of computational technology at the time. Since then, computer technology has been improving, specifically the speed of numerical computations. Accordingly, the flexibility studies have also been expanding to cover dynamic performances. These extensions are referred as the dynamic flexibility (Dimitriadis and Pistikopoulos, 1995; Mohideen *et al.*, 1996), or the controllability (Skogestad and Wolff, 1992) analyses. Further review of these developments is given in section 2.6.

Controllability is formally defined as 'the ability to achieve the desired process dynamic performance, within various limitations on process operations, despite the presence of disturbances and uncertainties, using the available manipulated and measured variables' (Skogestad and Wolff, 1992). In comparison to the flexibility, which merely sustains the effects of disturbances and uncertainties, the controllability appreciates the efforts of compensating the upsets and maintaining the original operating conditions.

The above definition can be translated into optimisation problems, in similar fashion to the flexibility analysis. Here, the process parameters are optimised to achieve the best process economic, ensuring that the process dynamics fit within the feasible operating space and performance specifications, despite the presence of set points changes, external disturbances, or uncertainties. The controller features, either ideal or realistic, can be augmented within the problem and assessed simultaneously. This approach clearly supports the idea of simultaneous process designs.

Another approach of controllability assessment is the analytical method. This approach determines the process controllability based on the attainability of a perfect inverse of

the process transfer function. The proponents argue that controllability must be an inherent feature of the process, and further control design must be based on this assessment. Clearly, this view is biased towards the sequential design procedure. Further reviews of these methods are given in section 2.7.

The Stability is defined as 'the ability to maintain bounded process dynamics in presence of bounded excitations'. Based on this definition, stability can be considered as a specific case in controllability analysis, where the bounded process dynamics become the constraints of the corresponding optimisation problem. Hence, while the process economy is the shared property of both controllability and flexibility, the stability is the specific aspect of controllability (Bahri, 1996).

Overall, both flexibility and controllability analyses cover major operability aspects. These analyses have been implemented using either the optimisation or the analytic methods. The features of optimisation methods in supporting flexibility and controllability are presented in the next section.

2.5 Optimisation in Chemical Process Designs

Optimisation methods have been the major component of the flexibility studies, as well as numerous controllability analyses, in the open literature. In this study, the optimisation method is considered as a medium to capture and analyse the operability features of process design and synthesis. Accordingly, this section provides an insight towards the complexity of the process design problems and the suitability of the available optimisation strategies.

2.5.1 Generic Formulation of Optimisation Problems

Optimisation has been the major decision making tool in various fields, including economy, business, engineering and science. The reason for its popularity is the capacity of formulating a wide range of problems in a concise manner, using the combination of process objectives and constraints. The objectives may include the

minimum economic costs or the maximum of certain performance indices. Meanwhile, the constraints represent the process model and the operational specifications. The goal of the optimisation is to find the process variables within the specification that return the best objective or the best trade-off between several objectives. A generic optimisation problem can be formulated as follows:

$$\begin{aligned}
&\max_{\bar{\mathbf{z}}} \sum_{\mathbf{q}} \boldsymbol{\Phi}_{\mathbf{q}}(\bar{\mathbf{z}}, \bar{\mathbf{w}}, \boldsymbol{\theta}, t) \qquad \mathbf{q} \in \mathbf{O} \\
&\text{s.t.} \quad \mathbf{h}_{\mathbf{i}}(\bar{\mathbf{z}}, \bar{\mathbf{w}}, \boldsymbol{\theta}, t) = \mathbf{0} \qquad \mathbf{i} \in \mathbf{E} \\
&\qquad \mathbf{g}_{\mathbf{j}}(\bar{\mathbf{z}}, \bar{\mathbf{w}}, \boldsymbol{\theta}, t) \le \mathbf{0} \qquad \mathbf{j} \in \mathbf{I} \\
&\qquad \bar{\mathbf{z}} = \{\mathbf{z}, \mathbf{y}\} \\
&\qquad \bar{\mathbf{w}} = \{\mathbf{x}, \mathbf{u}, \mathbf{w}\} \\
&\qquad \bar{\mathbf{z}} \in \bar{\mathbf{Z}} = \{\bar{\mathbf{z}} \mid \bar{\mathbf{z}}^{\mathbf{L}} \le \bar{\mathbf{z}} \le \bar{\mathbf{z}}^{\mathbf{U}}\} \\
&\qquad \bar{\mathbf{w}} \in \bar{\mathbf{W}} = (\bar{\mathbf{w}} \mid \bar{\mathbf{w}}^{\mathbf{L}} \le \bar{\mathbf{w}} \le \bar{\mathbf{w}}^{\mathbf{U}}) \\
&\qquad \boldsymbol{\theta} \in \boldsymbol{\Theta} \\
&\qquad t \in \{t \mid \mathbf{0} \le t \le t_f\}
\end{aligned} \tag{2.9}$$

Here, $\boldsymbol{\Phi}$ are the objectives, as functions of independent design variables $\bar{\mathbf{z}}$, dependent variables $\bar{\mathbf{w}}$, independent parameters $\boldsymbol{\theta}$ and time **t**. The variables can be specified further. For instance, $\bar{\mathbf{z}}$ may include the continuous design variables **z** and discrete design variables **y**. The dependent variables $\bar{\mathbf{w}}$ may include the state, the manipulated and the output variables, **x, u**, and **w** respectively. The independent parameters $\boldsymbol{\theta}$ and time **t** are not optimised within the problem, but significantly affect the optimisation results. There can be **O** objectives. The process model is represented by **E** equalities in the set $\mathbf{h}_i$, and the process specification by **I** inequalities in the set $\mathbf{g}_j$.

The variables are feasible if they fall within the respective allowable set and satisfy the constraints in (2.9). The vector $\bar{\mathbf{z}}^*$ is optimal if it is feasible, and the corresponding objective value is the maximum of all feasible solutions as given in (2.10).

$$\sum_{\mathbf{q}} \boldsymbol{\Phi}_{\mathbf{q}}(\bar{\mathbf{z}}^*, \bar{\mathbf{w}}) \ge \sum_{\mathbf{q}} \boldsymbol{\Phi}_{\mathbf{q}}(\bar{\mathbf{z}}, \bar{\mathbf{w}}) \tag{2.10}$$

Here, the sense of optimisation is presented as maximisation. However, it could just as well be minimisation, by reversing the sign in (2.10) (Bryson, 1999).

2.5.2 Optimisation Tree

The tree of optimisation problems is illustrated in Figure 2.2. The linear unconstrained optimisation problem is the simplest branch of the optimisation tree. The constrained optimisation algorithms are the generally extension of this problem, where the constraints may take the form of any or some combinations of bound constraints, nonlinear constraints, discrete variables, or stochastic variables and functions.

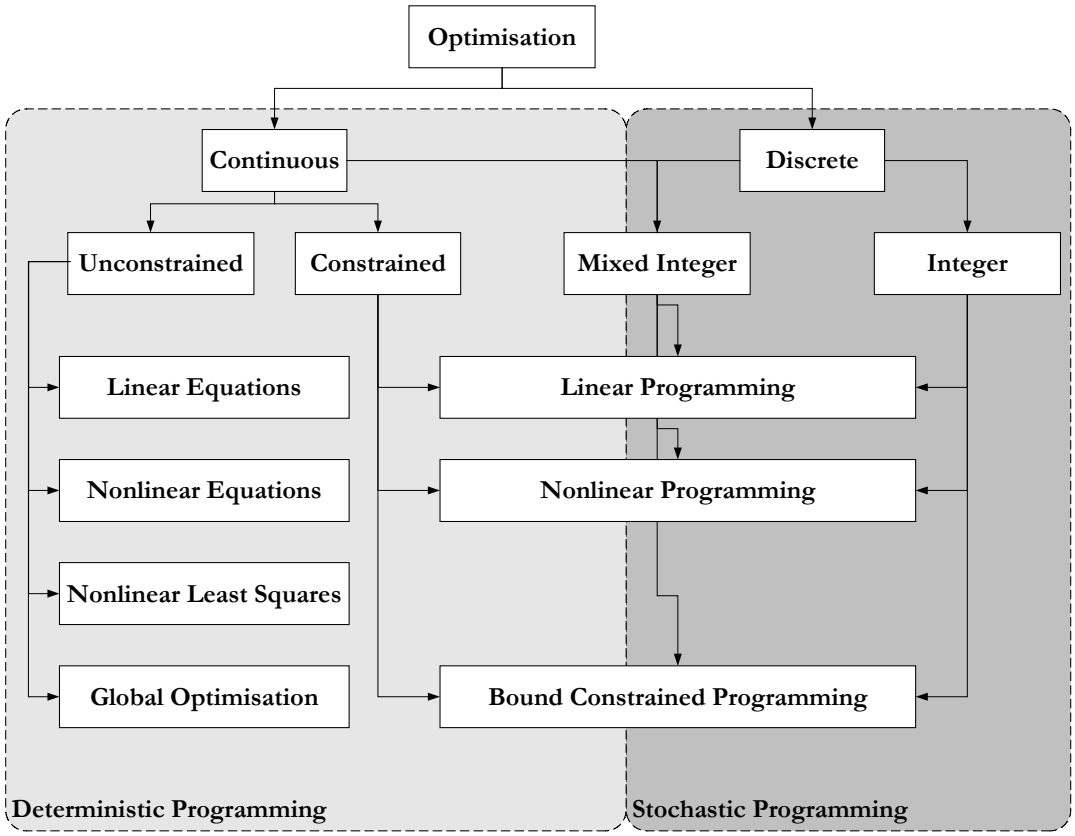


Figure 2.2 Optimisation Spectrum

Typical chemical process design problems are located at the various branches of the optimisation tree. The process models are generally nonlinear, and these transform the design problem into the Nonlinear Programming (NLP) problem. Typically, the operational conditions are constrained within the upper and lower limits, leading to the bound optimisation problem. The variables involved may not be continuous (real) within the given range of values, but consist of several discrete values. This case is the Discrete Programming problem. When the values are all integers, it is an Integer Programming (IP) problem, while the general cases would be mixed real and integer case, hence a Mixed Integer Programming (MIP) problem (Bryson, 1999).

Some process variables may not be known accurately for various reasons. For instance, some of the parameters are too expensive to measure, the exact dynamic model is too difficult to derive, or some values are predicted for a future time, none of which can be known with certainty. It can be said that these cases contain uncertainties θ, which are assumed as random variables or functions. The respective optimisation problem is Stochastic Programming, which is contrasted with Deterministic Programming, where all variables are known with certain accuracy.

Finally, a number of different objectives may be optimised simultaneously, resulting in the Multi-objective, as opposed to the Single-objective programming problem. For instance, the controllability assessment is a multi-objective problem, which maximises both the profit and dynamic performances. It is possible that these objectives are conflicting, leading to a trade-off problem. The common approach is to reformulate the problem into a single-objective one. This can be achieved by forming a weighted combination of the different objectives. Alternatively, one objective is optimised, and the rest are transformed into weighted constraints (ε–constraint method) (Schweiger and Floudas, 1997).

2.5.3 Optimisation Framework

The nature of extending the established optimum solutions to address a specific problem raises the notion of an optimisation framework. Within this framework, the

problem is broken down and reorganised into simpler problems that can be solved efficiently, using readily available solvers. Therefore, the approach of solving the multi-objective problem as explained above can be considered as an optimisation framework, which solves the weighted-combination of single objective problems. With similar reasoning, a stochastic problem is considered as a framework consisting of multiple deterministic problems; and the mixed-integer problem is a framework with interconnected nodes, where each node applies tight constrain on one integer variable.

These approaches are common within the flexibility and controllability studies. The specific implementation of these fields in chemical process design is discussed in sections 2.6-2.8.

2.5.4 Translation of Chemical Process Designs into Optimisation Problems

The advances of computational technology have been the backbone of optimisation frameworks to solve process design problems. Nevertheless, a meaningful solution can only be deduced from a proper knowledge of the process, including critical analysis of the process and realistic process objectives based on the past experiences or 'engineering judgment' (Bahri, 1996). The analysis should determine the scope and quality of process design formulation. Specifically, this is translated into an attainable design objective, a proper process model, feasible operational ranges, selection and characterisation of process variables, as well as the consideration and characterisation of disturbances and uncertainties. The following descriptions discuss these translations in reference to the generic optimisation problem given in (2.8).

The objectives of process control design Φ may vary from profit, operational cost, down time, product quality, energy consumption, and human resources. The challenges are to determine the focus of the problem and to find the balance among any conflicting objectives. For example, reducing fuel consumption may save significant cost, but it may reduce the product purity, thus affecting the profit and further

processing. Determining the balance between these objectives depends on the engineering judgments, and leads to a multi-objective formulation.

The process constraints include the model and operational specifications. The process model $\mathbf{h}_i$, is typically derived from the mass and energy balances. The derivation should represent a wide range of process dynamics, but be simple enough to solve within a reasonable time. A linearised process model is often used, because the corresponding control design problem with a quadratic objective can be transformed into linear Riccatti matrix equation that can be solved efficiently (Anderson and Moore, 1990). However, this is only valid in small portions of the desired operational ranges, due to the local linearisation around an operating point. On the other hand, nonlinear model may cover wider operational range, but expensive to solve. Furthermore, the model may not be convex, and therefore the solution may not be unique (Marlin, 1995). The selection between these models would depend on the available computational tools.

The process specification $\mathbf{g}_j$ may involve the operational safety limits, environmental regulations and desired operational performances. Typically, the optimum operating condition is located within the subset of all constraints, namely the active constraints. As the process under study becomes more complex and the number of interacting variables increases, the feasible operational region defined by these constraints become more stringent. Therefore, the constraints should be assigned with due care, to avoid the conflicting constraints that give no feasible solutions.

The process designs generally involve both continuous and discrete variables. The continuous variables (**z**, **x**, **u**, **w**) represent the equipment sizes, the properties of process streams, such as flow rates and temperatures, which are continuous over time and variables spaces. The discrete variable **y** represents the existence of equipment, process streams, or logical relationships among process variables. These can also be used for defining different sets of process regimes. Mostly these are true – false decisions, which is represented by binary values [0, 1].

There are optimisation methods that depend on the gradients of the objective functions over the variable space to find the optimum points (Bryson, 1999). These solvers are effective for continuous problems, and fast convergence is guaranteed for convex cases. The discrete problems, on the other hand, cannot benefit from gradient information due to their discontinuous nature. Instead, these problems give rise to a search tree that grows exponentially according to the number of discrete variables and the discrete decisions within each variable. For example, if there are n_y discrete variables and each variable has P decisions, then there is a maximum of n_y^P branches to be optimised separately before the best solution is found. In general, the discrete problems are more expensive to solve than the continuous ones. Therefore, the discrete decisions should be formulated in such a way that the search is minimised.

The disturbances and uncertainties θ are the important considerations in process design. These loads perturb the process from the optimum steady state operating conditions, which may cause the constraint violations. To keep the perturbed process feasible, the operating conditions should be moved further inside the feasible region. Since these uncertain parameters are only known approximately, the corresponding change in the operating conditions must be treated judiciously as the optimistic expectation. The smaller the effects of disturbances and uncertainties, the closer the process can be operated to the previously steady state economic optimum. The effects are not only determined by their own characterisations, such as the functions, magnitudes or frequency of occurrences, but also on the process inherent ability to reject them, and the implemented controller (Bahri, 1996). The relationships between these aspects are prominent in operability analysis, especially in flexibility and controllability studies. These studies are reviewed in the following sections.

2.6 Flexibility Analysis

The process flexibility, or feasibility, indicates the capability to maintain the process within feasible operational region in the presence of uncertainties (Grossmann *et al.*,

1983). This nature is incorporated at the design stage of the process, where it is ensured that the economic specification is satisfied for a given range parameter values.

The analysis generally involves two complementary tasks, which are the calculation of flexibility index and the flexibility test. The flexibility index quantifies the feasible operation of a given design, in the presence of the assumed ranges of disturbances and uncertainties. Reciprocally, the flexibility test is the assessment of the capability to operate feasibly against the known index. Either task takes the form of the following optimisation problem:

$$\max_{\bar{z}} \Phi(\bar{z}, \bar{w}, \theta, t) \tag{2.11}$$

$$\begin{aligned}
\text{s.t.} \quad & h_i(\bar{z}, \bar{w}, \theta, t) = 0 && i \in E \\
& g_j(\bar{z}, \bar{w}, \theta, t) \le 0 && j \in I \\
& \bar{z} = \{z, y\} \\
& \bar{w} = \{x, u, w\} \\
& \bar{z} \in \bar{Z} = \{\bar{z} \mid \bar{z}^L \le \bar{z} \le \bar{z}^U\} \\
& \bar{w} \in \bar{W} = (\bar{w} \mid \bar{w}^L \le \bar{w} \le \bar{w}^U) \\
& \theta \in \Theta \\
& t \in \{t \mid 0 \le t \le t_f\}
\end{aligned}$$

Here, Φ is the objective, a function of the continuous design variables $\mathbf{z}$ and discrete design variables $\mathbf{y}$. The dependent variables $\bar{w}$ may include the state, the manipulated and the output variables, $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{w}$ respectively. The process is affected by uncertain parameters θ over time t. The process model is represented by $\mathbf{E}$ equalities in the set $\mathbf{h}_i$, and the process specification by $\mathbf{I}$ inequalities in the set $\mathbf{g}_j$.

Generally, only a limited information about the uncertainties θ can be reasonably gathered, such as the highest and lowest expected values, sample averages and variances, or the approximated profiles for a certain time horizon. Due to this approximate knowledge, θ is assumed to have infinite possible values in addition to the given information. This assumption transforms the flexibility problem into a stochastic semi-infinite optimisation problem.

This semi-infinite problem can be handled using either the deterministic, or the stochastic approaches. The deterministic method assumes that the effect of the whole set $\theta \in \Theta$ is represented by the critical values θ^k. The values can be found using the vertex enumeration procedure, the active constraints strategy, or the parametric method. Meanwhile, the stochastic method explicitly assumes some kind of a probabilistic property over the set $\theta \in \Theta$, such as Gaussian probability distributions. The method approximates the expected values over this distribution using a weighted-multi-period optimisation of the uncertainty samples $\theta_q \in \Theta$. These approaches are reviewed further in section 2.6.1-2.6.2.

The scopes of the flexibility studies reviewed in the following section are located within the spectrum illustrated in Figure 2.2. Over decades, the developed solutions have been following the growth of the optimisation tree described in section 2.5.2 with support from the advances of the computation technology. The outcome is that the complexity has been increasing from the linear, steady state deterministic cases to the nonlinear, dynamic, stochastic counterparts.

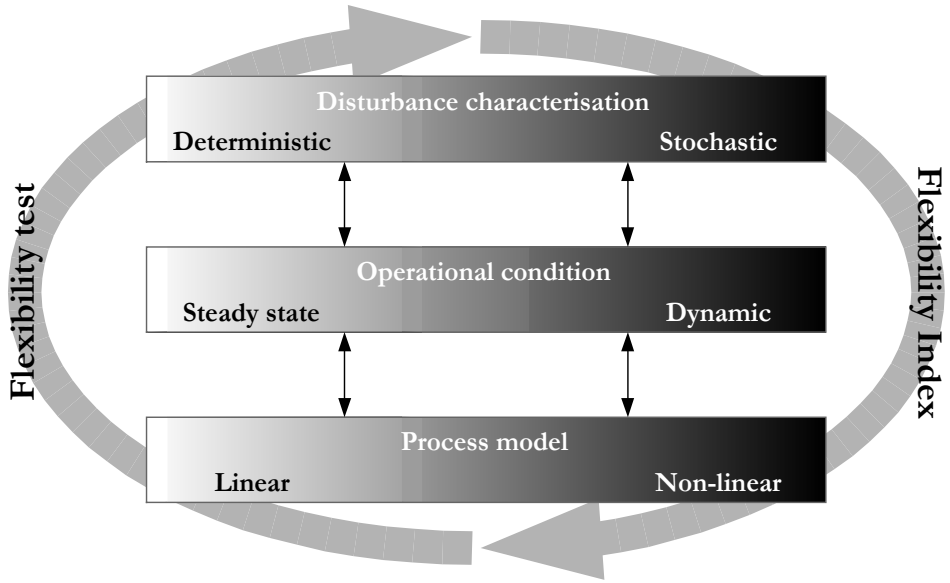


Figure 2.3 The Scopes of the Flexibility Studies

2.6.1 Deterministic Flexibility Analysis

The early flexibility studies were focused on the steady state process model as follows:

$$\mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le \mathbf{0} \quad \mathbf{i} \in \mathbf{E} \tag{2.12}$$

$$\mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{x},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le \mathbf{0} \quad \mathbf{j} \in \mathbf{I} \tag{2.13}$$

Here, $\mathbf{h}_i$ represents the set of model equations and $\mathbf{g}_j$ represents the set of inequalities defining the feasible operational region. The variables $\mathbf{z}$, $\mathbf{y}$, $\mathbf{u}$, $\mathbf{w}$, and $\boldsymbol{\theta}$ are the vectors of continuous design variables, discrete design variables, manipulated variables, output variables and uncertain parameters, respectively.

Halemane and Grossmann (1983) set up a feasibility test problem to determine the set of design values that maintain a given linear process within the feasible operating region, despite the occurrence of the set of uncertain parameters. The uncertain parameters are given in their nominal values $\boldsymbol{\theta}^N$, the expected deviations in the positive and negative directions, $\Delta\boldsymbol{\theta}^+$ and $\Delta\boldsymbol{\theta}^-$ respectively, and correlation $\mathbf{r}(\boldsymbol{\theta})\le\mathbf{0}$ as follows:.

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta} \mid \boldsymbol{\theta}^N - \Delta\boldsymbol{\theta}^- \le \boldsymbol{\theta} \le \boldsymbol{\theta}^N + \Delta\boldsymbol{\theta}^+, \mathbf{r}(\boldsymbol{\theta}) \le 0\} \tag{2.14}$$

The corresponding feasibility test is formulated in the following optimisation problem:

$$\psi(\mathbf{z},\mathbf{y},\boldsymbol{\theta}) = \min_{\mathbf{z}\in\mathbf{Z}} \upsilon \tag{2.15}$$

$$\begin{aligned} \mathbf{s.t.} \quad & \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \mathbf{0} \quad \mathbf{i} \in \mathbf{E} \\ & \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le \upsilon \quad \mathbf{j} \in \mathbf{I} \end{aligned}$$

Here, υ is the distance to the active constraint, ψ is the constraint projection function, where $\psi\le0$ indicates feasibility and $\psi>0$ the infeasibility of process design upon a fixed value of $\boldsymbol{\theta}$. In terms of ψ, the feasibility is established in the following formulation:

$$\chi(\mathbf{z},\mathbf{y}) = \max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \psi(\mathbf{z},\mathbf{y},\boldsymbol{\theta}) \tag{2.16}$$

Accordingly, $\chi\le0$ indicates the feasibility of design variables $\mathbf{z}$ and $\mathbf{y}$ for $\boldsymbol{\theta}\in\boldsymbol{\Theta}$. The term $\chi>0$ indicates otherwise, where (2.16) gives the critical points $\boldsymbol{\theta}^*$ causing the largest constraint violations.

The formulations (2.15) and (2.16) simply indicate whether the design is feasible. The feasibility is quantified by the Flexibility Index (FI), which is the largest scaled deviation δ of the expected deviations that can be handled by the design within a feasible operation. The formulation is as follows:

$$\begin{aligned} &\mathrm{FI}(\mathbf{z},\mathbf{y}) = \max \delta && (2.17)\\ &\textbf{s.t.} \quad \max_{\theta \in \boldsymbol{\Theta}(\delta)} \psi(\mathbf{z},\mathbf{y},\boldsymbol{\theta}) \le 0 \\ &\qquad \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = 0 && \mathbf{i} \in \mathbf{E} \\ &\qquad \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le 0 && \mathbf{j} \in \mathbf{I} \\ &\qquad \boldsymbol{\Theta}(\delta) = \{\boldsymbol{\theta} \mid \boldsymbol{\theta}^N - \Delta\boldsymbol{\theta}^- \le \boldsymbol{\theta} \le \boldsymbol{\theta}^N - \Delta\boldsymbol{\theta}^+, \mathbf{r}(\boldsymbol{\theta}) \le 0\} \\ &\qquad \delta \ge 0 \end{aligned}$$

A value of FI≥1 indicates that the design can handle even wider ranges of uncertainty than the given $\boldsymbol{\Theta}$. On the contrary, FI<1 indicates that only a fraction of the expected deviations can be tolerated.

The flexibility analysis given in (2.16)-(2.17) assumes that both the model $\mathbf{h}_i$ and the constraints $\mathbf{g}_j$ are linear, and that the uncertainties are uncorrelated ($\mathbf{r}(\boldsymbol{\theta}) \le \mathbf{0}$). Therefore, ψ is monotonically increasing or decreasing over the space $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, and the critical point $\boldsymbol{\theta}^{k*}$ is located in one of the vertices of the multi-dimensional rectangle $\boldsymbol{\Theta}$. Based on this condition, the feasibility analysis (2.16)-(2.17) can be reduced to the evaluation of each vertices of $\boldsymbol{\Theta}$, or the vertex enumeration scheme (Swaney and Grossman, 1985a).

The formulation of vertex enumeration for the feasibility test is as follows:

$$\begin{aligned} &\chi(\mathbf{z},\mathbf{y}) = \max_{\boldsymbol{\theta}^k \in \boldsymbol{\Theta}} \upsilon^k && (2.18)\\ &\textbf{s.t.} \quad \upsilon = \min_{\upsilon, z} \upsilon \\ &\qquad \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}^k) = \mathbf{0} \quad \mathbf{i} \in \mathbf{E} \\ &\qquad \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}^k) \le \upsilon \quad \mathbf{j} \in \mathbf{I} \end{aligned}$$

Here, $\boldsymbol{\theta}^k \in \boldsymbol{\Theta}$ is the set of the vertices in the corresponding space.

Similarly, the vertex enumeration formulation of the feasibility index is as follows:

$$\mathrm{FI}(\mathbf{z},\mathbf{y}) = \max_{\theta^k \in \boldsymbol{\Theta}} \delta^k \tag{2.19}$$

$$\begin{aligned}
\textbf{s.t.}\quad & \delta^k = \max_{\delta,\mathbf{z}} \delta \\
& \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\theta^k) = \mathbf{0} && i \in \mathbf{E} \\
& \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\theta^k) \le \upsilon && j \in \mathbf{I} \\
& \boldsymbol{\Theta}(\delta) = \{\theta \mid \theta^N - \Delta\theta^- \le \theta \le \theta^N - \Delta\theta^+, \mathbf{r}(\theta) \le 0\} \\
& \delta \ge 0
\end{aligned}$$

Both (2.18) and (2.19) involve $2^{n\theta}$ optimisation problems, where n_θ is the number of the uncertainty parameters. The size of this problem grows exponentially with n_θ. In order to reduce the number of evaluations, the implicit vertex enumeration schemes have been proposed (Kabatek and Swaney, 1992; Swaney and Grossman, 1985b). In parallel studies, Bandoni and Romagnoli (1989) proposed an iterative cutting plane algorithm to solve the flexibility problems using the linear programming method. They extended the procedure for linear problems under uncertainty constrained by inequalities (Friedman and Reklaitis, 1975).

As an alternative to the vertex enumeration method, Grossmann and Floudas (1987) developed a mixed-integer formulation for the flexibility test and index problems using the active constraints strategy. The basis of this strategy is that the feasibility function ψ is a piecewise continuous function for a particular design and structure, and the segments are characterised by a set of active inequality constraints $\mathbf{I_A^k}$. This active set is the subset of constraints at (2.15) that have non-zero multipliers and satisfy the Kuhn-Tucker conditions as follows:

$$\sum_{j \in \mathbf{I}_A^k} \lambda_j^k = 1 \tag{2.20}$$

$$\sum_{j \in \mathbf{I}_A^k} \lambda_j^k \frac{\partial \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\theta^k)}{\partial \mathbf{z}} = 0 \tag{2.21}$$

$$\psi^k = \sum_{j \in \mathbf{I}_A^k} \lambda_j^k \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\theta^k) \tag{2.22}$$

The number of the active constraints $\mathbf{I_A^k}$ is equal to the number of design variable plus one. It is assumed that the gradients of the constraints are linearly independent. The information is used to simultaneously find the subset during the computation of the feasibility test and index. This leads to a mixed-integer optimisation problem (MIP).

The corresponding flexibility test based on the active constraint strategy is as follows:

$$\chi(\mathbf{z},\mathbf{y}) = \max_{\theta\in\Theta} \upsilon \tag{2.23}$$

$$\begin{aligned}
\textbf{s.t.}\quad & \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \mathbf{0} && \mathbf{i}\in\mathbf{E}\\
& \mathbf{s}_j + \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \upsilon && \mathbf{j}\in\mathbf{I}\\
& \sum_{j\in\mathbf{I}} \lambda_j = 1 && \mathbf{j}\in\mathbf{I}\\
& \sum_{j\in\mathbf{I}} \lambda_j \frac{\partial \mathbf{g}_j}{\partial \mathbf{z}} = 0 &&\\
& \lambda_j - \mathbf{y}_j \le 1 && \mathbf{j}\in\mathbf{I}\\
& \mathbf{s}_j - \mathbf{M}(1-\mathbf{y}_j) \le 1 && \mathbf{j}\in\mathbf{I}\\
& \sum_{j\in\mathbf{I}} \mathbf{y}_j = n_z + 1 && \mathbf{j}\in\mathbf{I}\\
& \boldsymbol{\Theta}(\delta) = \{\boldsymbol{\theta} \mid \boldsymbol{\theta}^N - \boldsymbol{\Delta\theta}^- \le \boldsymbol{\theta} \le \boldsymbol{\theta}^N - \boldsymbol{\Delta\theta}^+, \mathbf{r}(\boldsymbol{\theta}) \le 0\} &&\\
& \mathbf{y}_j = \{0,1\};\quad \mathbf{s}_j \ge 0 && \mathbf{j}\in\mathbf{I}
\end{aligned}$$

Here, the expression $\mathbf{s_j - M(1-y_j)\le 1}$ is utilised to find the active constraints $\mathbf{I_j}$, which is indicated by the slack variable $\mathbf{s_j}$=0 and $\mathbf{y_j}$=1. The constant **M** can be any big real value; most often, **M** equals to the valid upper bound of the constraint. This way, the corresponding constraint is activated or deactivated the according to the value of **y**, 1 or 0, respectively. This strategy is known as the Big-M approach.

The corresponding flexibility test using this strategy is as follows:

(2.24)

$$\begin{aligned}
& \mathrm{FI} = \min_{\theta\in\Theta} \delta &&\\
\textbf{s.t.}\quad & \mathbf{h}_i(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \mathbf{0} && \mathbf{i}\in\mathbf{E}\\
& \mathbf{s}_j + \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \upsilon && \mathbf{j}\in\mathbf{I}\\
& \sum_{j\in\mathbf{I}} \lambda_j = 1 && \mathbf{j}\in\mathbf{I}
\end{aligned}$$

$$\sum_{j\in \mathbf{I}} \lambda_j \frac{\partial \mathbf{g}_j}{\partial \mathbf{z}} = 0$$

$$\lambda_j - y_j \le 1 \qquad j \in \mathbf{I}$$

$$s_j - \mathbf{M}(1 - y_j) \le 1 \qquad j \in \mathbf{I}$$

$$\sum_{j\in \mathbf{I}} y_j = n_z + 1 \qquad j \in \mathbf{I}$$

$$\boldsymbol{\Theta}(\delta) = \{\theta \mid \theta^N - \Delta\theta^- \le \theta \le \theta^N - \Delta\theta^+, \mathbf{r}(\theta) \le 0\}$$

$$y_j = \{0, 1\}; \quad s_j \ge 0 \qquad j \in \mathbf{I}$$

While this approach is applicable to both linear and nonlinear problems, it requires the solution of mixed integer linear and nonlinear problems (MILP and MINLP, respectively). The MILP problems can be solved using the standard Branch and Bound (BB) method; while the MINLP problems can be solved using either the Generalised Bender Decomposition (GBD) method (Geoffrion, 1972), or the Outer Approximation (OA) method. The variants of the last method include the Outer Approximation with Equation Relaxation (OA/ER) and Outer Approximation with Equation Relaxation and Augmented Penalty (OA/ER/AP) methods (Duran and Grossmann, 1986; Kocis and Grossmann, 1987; Viswanathan and Grossmann, 1990). Further discussion about various MILP and MINLP solvers are given in Chapter 5, which presents the development of the Branch and Bound solution for nonlinear problems.

The global solution of the flexibility problems using the active constraint strategy is presented in Floudas *et al.* (1999; 2001), where the nonlinear constraints $\mathbf{g}_j$ are underestimated using several convex functions. The resulting MINLP is solved using the BB or the GBD methods. Recently, Raspanti *et al.* (2000) over-estimated the active constraints into a single piecewise function using the constraint aggregation method.

Pistikopoulos and Grossmann (1988a; 1988b) defined the feasibility function ψ as an analytical function of uncertain parameter θ and design variable $\mathbf{z}$. The approach has been extended to special nonlinear cases (Pistikopoulos and Grossmann, 1988c; d), then further evolved into the Parametric Programming Framework (Acevedo and

Pistikopoulos, 1996; Bansal *et al.*, 2000; 2002a; b; Varvarezos *et al.*, 1995; Vassiliadis *et al.*, 1994). These frameworks provide the profile of the process objective at a targeted flexibility index, over the spaces of design variables $\mathbf{z}$ and uncorrelated uncertainty parameters $\boldsymbol{\theta}$.

The parametric programming approach uses the general formulation as follows:

$$\begin{aligned} &\psi(\mathbf{z},\mathbf{y}) = \max_{\theta\in\Theta} \upsilon && (2.25)\\ &\textbf{s.t.} \quad \mathbf{H}_z\mathbf{z} + \mathbf{H}_y\mathbf{y} + \mathbf{H}_u\mathbf{u} + \mathbf{H}_w\mathbf{w} + \mathbf{H}_\theta\boldsymbol{\theta} + \mathbf{h}_c = 0 \\ &\qquad \mathbf{g}_j(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \upsilon\mathbf{e} \qquad\qquad \mathbf{j}\in\mathbf{I} \end{aligned}$$

Here, $\mathbf{H}_l$, $\mathbf{l = z, y, u, w, \theta}$ are constant matrices, $\mathbf{h}_c$ is a vector of constants, $\mathbf{e}$ is a unit vector and $\mathbf{g}_j$ is convex. Assuming the integer variables $\mathbf{y}$ are relaxed, (2.25) corresponds to a convex multi-parametric nonlinear programming, where the feasibility function ψ is a linear function of the involved parameters (Dua and Pistikopoulos, 1999).

Using these properties, the linear profile of ψ can be approximated. Firstly, an outer approximation of (2.25) is created by linearising the constraints $\mathbf{g}_j$ at an initial feasible point. This yields a multi-parametric linear program, which is solved to give a set of linear, underestimating parametric expressions for ψ and a corresponding set of linear inequalities defining the optimal region. The maximum discrepancy between the real objective and the linear approximation would be located at one of the vertices of this region. If the discrepancies are greater than a user specified tolerance ε, then the corresponding vertex becomes an additional point of linearisation for the additional approximation region. The procedure is iterated until all of the approximated regions are within the desired tolerance (Dua and Pistikopoulos, 1999).

Since the size of MILP/MINLP problems in a parametric approach is generally large, Varvarezos *et al.* (1995) used the sensitivity information of the uncertain parameters to screen the non-redundant set of active constraints of linear cases. Based from this approach, the parametric algorithm is gradually established through the works of

Vassiliadis *et al.* (1994), Acevedo and Pistikopoulos (1996), Dua and Pistikopoulos (1999) and Dua *et al.* (2002). Bansal *et al.* (2000; 2002a; 2002b) combined the algorithms into a unified framework for the cases with uncorrelated uncertainties. Rooney and Biegler (2003) recently applied this approach for a steady state process with correlated uncertainties, which are defined within a confidence region derived from a likelihood ratio test.

The Resilience Index (RI) (Saboo *et al.*, 1985) is one index with similar concept to the Flexibility Index. It was introduced as a steady state measurement of the capability of a heat exchanger network (HEN) to deal with the changes in the inlet and target temperatures. The RI is defined as the maximum disturbance load that a network can allow without becoming infeasible as follows:

$$\mathbf{RI} = \max_{\mathbf{z},\boldsymbol{\theta}} \sum_{i} |\theta_i| \tag{2.26}$$

This problem can be generalised into any nonlinear system with disturbances $\boldsymbol{\theta}$ and system variables $\mathbf{z}$. Unlike FI, the RI cannot describe an asymmetric flexibility, but the computational cost is significantly cheaper.

These deterministic flexibility studies have been searching for a single critical realisation of uncertainty to determine the flexibility of the process. In the next section, stochastic flexibility analyses involving various types of uncertainties are reviewed.

2.6.2 Stochastic Flexibility Analysis

The flexibility analysis using the stochastic approach explicitly assumes that the disturbances and uncertainties $\boldsymbol{\theta}$ are characterised by continuous probability density functions (PDF). The associated optimisation problems therefore become stochastic programming. In this problem, either the objective functions or the constraints may be expressed in terms of some probabilistic representation, such as expected value or variances.

The general formulation of stochastic problem is as follows (Kim and Diwekar, 2002):

$$\min_{\mathbf{z}} \; P_1\left[\Phi(\bar{\mathbf{z}}, \boldsymbol{\theta})\right] \tag{2.27}$$

$$\text{s.t.} \quad \mathbf{P_2}\left[\mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta})\right] = 0 \quad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{P_3}\left[\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta})\right] \le 0 \quad \mathbf{j} \in \mathbf{I}$$

$$\bar{\mathbf{z}} \in \bar{\mathbf{Z}}$$

$$\boldsymbol{\theta} \in \boldsymbol{\Theta}$$

Here, $\bar{\mathbf{z}}$ is a vector of decision variables and $\boldsymbol{\theta}$ is a vector of disturbances and uncertainties. The objective function $P_1[\Phi]$ is optimised subject to the model $\mathbf{P_2}[\mathbf{h_i}]$ and the operational constraints $\mathbf{P_3}[\mathbf{g_j}]$, where any or all of them are probabilistic functions. For example, if P_1 is the expected value, the above optimisation problem becomes:

$$\min_{\mathbf{z}} \; \mathbf{P_1}\left[\Phi(\bar{\mathbf{z}}, \boldsymbol{\theta})\right] = \min_{\mathbf{z}} \; E_{\theta}\left[\Phi(\bar{\mathbf{z}}, \boldsymbol{\theta})\right] \tag{2.28}$$

Here, E_θ is the expected value over $\boldsymbol{\theta}$.

Grossmann *et al.* (1983) and Halemane and Grossman (1983) employed the above formulation to expand the concept of flexibility index into the optimal design with a fixed degree, or with an optimal degree of flexibility. The objective of the fixed flexibility degree problem is to find the optimum expected cost, E_θ, of the optimal design that can be operated feasibly over the explicit probability set $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (equivalent to FI = 1) as follows:

$$\min_{\mathbf{z}\in\mathbf{Z},\mathbf{y}\in\mathbf{Y}} E_{\theta}\left[\min_{\mathbf{u}} \boldsymbol{\Phi}(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \,|\, \mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \mathbf{0}, \mathbf{g_j}(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le 0\right] \tag{2.29}$$

$$\mathbf{s.t.} \quad \max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \min_{\mathbf{u}} \max_{\mathbf{j}\in\mathbf{I}} \mathbf{g_j}(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) \le 0 \quad \mathbf{j} \in \mathbf{I}$$

$$\mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{u},\mathbf{w},\boldsymbol{\theta}) = \mathbf{0} \qquad \mathbf{i} \in \mathbf{E}$$

Here, the design variables $\bar{\mathbf{z}}$ is broken down into the fixed design variables **z**, **y**, and the manipulated variable **u**. The fixed design variables determine the optimum design based on the parameter realisations, while the manipulated variable compensates the effects of $\boldsymbol{\theta}$.

The optimal design with optimal degree of flexibility is the generalisation of the above formulation. This time, the objectives are to simultaneously minimise the cost and maximise the flexibility index, while ensuring the feasibility operation over the set $\theta \in \Theta$. This can be formulated as follows:

$$\begin{aligned} &\min_{z\in Z, y\in Y} E_\theta\left[\min_{u} \Phi(z,y,u,w,\theta) \mid h_i(z,y,u,w,\theta)=0, g_j(z,y,u,w,\theta)\le 0\right] \\ &\max_{u,\delta} \delta \\ &\text{s.t.} \quad \max_{\theta\in\Theta}\min_{u}\max_{j\in I} g_j(z,y,u,w,\theta)\le 0 \quad j\in I \\ &\qquad h_i(z,y,u,w,\theta)=0 \qquad\qquad i\in E \end{aligned} \tag{2.30}$$

The above formulation is a conflicting multi-objective problem, where improving one objective will put pressure on the other. When $\delta = 1$, (2.30) is equivalent to (2.29).

For a system with continuous uncertain parameters described by a joint PDF, the Stochastic Flexibility (SF) is the measurement of the probability for a feasible operation of a given design (Pistikopoulos and Mazzucchi, 1990; Straub and Grossman, 1990). This index is formulated as follows:

$$SF(z,y) = P(\psi(z,y,u,w,\theta)\le 0) \tag{2.31}$$

When the uncertainties include a discrete probability, the Expected Stochastic Flexibility (ESF) (Pistikopoulos and Mazzucchi, 1990; Straub and Grossman, 1990) is defined as follows:

$$ESF(z) = \sum_{s=S_1}^{2n_s} SF(z,y^s)\,P(y^s) \tag{2.32}$$

Here, s is the index set for the process states, $P(y^s)$ is the discrete probability that the process is in state s and n_s is the number of possible states in the process.

Both SF and ESF concepts assume that the continuous uncertainties propagate through the process, causing the uncertainties in the output variables. Typically, the feasible output variables are defined within the constraints. Accordingly, the SF is

defined as the portion of the output distribution within the constraints, hence represent the probability of feasible operation.

However, it is difficult to propagate the distribution through the process and to ensure all constraints is accounted for. To solve this problem, Straub and Grossmann (1992) projected the constraints into the space $\theta \in \Theta$, and accordingly presented the model and the solution methods to evaluate the SF and ESF. This gives the design variables that optimise the indices under a parametrically varying cost constraint, and establishes the trade-off between the flexibility and the cost. Pistikopoulos and Mazzucchi (1990), on the other hand, propagated the uncertain parameter distribution in the space of ψ in combination with the parametric programming approach.

Grossmann and Sargent (1978) solved the fixed feasibility problem (2.29) for linear and steady state problem using the Gaussian quadrature formula. Here, the uncertainties can be characterised by any type of PDF, as long as they are uncorrelated. The basic idea of the Cartesian integration method is to approximate the multiple integral of the expected cost E_θ over the set $\theta \in \Theta$, through the Gaussian quadrature of n_θ-1 uncertain parameters. The set $\theta \in \Theta$ is discretised at the points determined by the quadrature formula, and the corresponding weighed cost is evaluated at each node. This feasibility problem is transformed into a multi-period optimisation. In this study, only the manipulated variables are used to compensate the uncertainty effects.

The major drawback of this method is that the number of the considered nodes increases exponentially with n_θ. To reduce the size of the problem, Pistikopoulos and Grossmann (1988b; 1988c; 1988d; 1988e) performed the sensitivity analysis to select and then use the subset of $\theta \in \Theta$ that significantly affected the cost. This modified method has been applied to solve the flexibility index problems (2.29)- (2.30) for design and retrofit of linear and nonlinear steady state cases. This include the design and operation of batch processes (Epperly *et al.*, 1997; Ierapetritou *et al.*, 1996; Ierapetritou and Pistikopoulos, 1996), and the calculation of the flexibility function of complex systems (Ostrovsky *et al.*, 1998).

Novak and Kravanja (1999) proposed another reduction technique called the Reduced Dimensional Stochastic (RDS) method. This method calculates the conditional expected objective functions using simultaneous optimisation in n_q quadrature points for each uncertain parameter while keeping other uncertain parameters at nominal values. Using linear programming, the excess conditional values are eliminated, yielding the basic vertices and the corresponding weighting values. The number of the basic vertices may be further reduced heuristically, for instance, by n_r. The maximum number of objective function evaluation using this procedure for n_θ number of uncertainties is $2^{n\theta-nr}+n_\theta+1$.

The procedure generates various sets of basic vertices and many sub optimal solutions. Therefore, Novak and Kravanja (2003) went further to produce only one central basic point. They utilised the objective values on n_q points of each uncertain parameter to build an approximate function using curve fitting. Then the extreme values of these functions are mapped back to the critical point of each uncertain parameter. Finally, the linear combination of these critical points determines the central basic point in uncertainty space to approximate the expected value of the objective function.

One alternative to the Gaussian quadrature method is applying the probabilistic samples of $\theta \in \Theta$. The examples include the usage of the Monte Carlo method, Latin Hypercube Sampling (LHS) (Diwekar and Rubin, 1991; Liu and Sahinidis, 1996) or the Hammersley Sequence Sampling (HSS) method (Bernardo and Saraiva, 1998; Kim and Diwekar, 2002). Here, the samples are produced from a pseudo-random number generator. This approach depends on the quality of the respective random number generators in producing the representative samples. The comparison of the quality of the existing sampling methods is given in Kim and Diwekar (2002). While the number of samples does not grow with n_θ, the number of samples required is arbitrarily large in order to guarantee the quality of the expected cost E_θ. Due to the size of the problem, the reported applications within this line of research are limited to steady state problems (Ahmed and Sahinidis, 1998; Bernardo and Saraiva, 1998; Samsatli *et al.*, 1998).

The common assumption behind the deterministic and stochastic flexibility studies reviewed so far has been that the processes operate at steady state for any value of the uncertain parameters. However, these analyses are of little help during the actual dynamic process operation, which may involve startups or shutdowns, transitions from one operating point to another, and disturbance rejections (Grossman and Morari, 1983).

The consideration of these dynamic properties in process marks the shift of the operability studies into the controllability analysis. Within this area, the optimisation-based flexibility problems are extended along with other controllability analysis approaches. These methods are reviewed in the following section.

2.7 Controllability Analysis

The Controllability concept has been used with two different interpretations in literature. The first is based on the goal-oriented view, while the other is based on the mathematical view. The discussion in this section focuses on the goal-oriented view, and is complemented with brief comment on the mathematical oriented definition.

The goal oriented view of controllability analysis can be traced back to Ziegler and Nichols (1943), who defined the controllability as 'the ability of the process to achieve and maintain the desired equilibrium value'. This definition is inline with the cited definitions by Skogestad and Wolff (1992) in section 2.4. This definition considers the controllability as a process property that indicates how easy it is to control the process to achieve the desired measurable performances. It also acknowledges that process control performances depend on the availability of both manipulated and measured variables. Accordingly, the concept is referred to as the 'input-output controllability'.

The concept is different from the 'state controllability' definition (Kalman, 1960). The latter says 'A state is termed controllable if for any initial state $x(0)=x_0$, any time t_1 and final state x_1, there exists an input u(t) such that $x(t_1)=x_1$'. This definition is based purely on a mathematical view. The dual is the 'state observability' concept, which is

defined as follows: 'A state is termed observable at some given t if knowledge of the input u(t) and output w(t) over a finite time $t_0<t<t$ completely determines x(t)'. The state controllability requires that the information of all state variables is somehow available. It also requires that the available inputs and measurements can reach all of the state variables.

The combination of the state controllability concept and the linear process produces various elegant and mathematically tractable formulations for process analyses. Its nice mathematical property has been widely used in linear controllability analyses for about four decades, some of which are reviewed in section 2.7.1. However, this concept does not indicate the ease of controlling a process, accommodate any specifications over the trajectory between the initial and final points, nor apply to general cases where not all states are controllable nor observable (Chenery, 1997). In this case, it is argued that the 'state controllability' is a specific case of the 'input-output controllability' concept.

With support from the advances of computational technology, nonlinear controllability analyses took shape in analytical and optimisation methods for the last two decades. The 'input-output controllability' concept has been gradually applied, starting from the open loop controllability cases. These are reviewed in section 2.7.2, followed with the discussion of integrated controllability assessment in process design in section 2.8.

2.7.1 Linear Controllability Analysis

The linear controllability analysis assumes linear process models and considers the mismatch between the model and the actual process as the part of process uncertainties.

The linearised process models are commonly derived from the local linearisation of the original explicit DAE and represented as follows:

$$\dot{\mathrm{x}}(\mathrm{t}) = \mathbf{A}\mathbf{x}(\mathrm{t}) + \mathbf{B}\mathbf{u}(\mathrm{t}) + \mathbf{F}\boldsymbol{\theta}(\mathrm{t}) \tag{2.33}$$

$$\mathrm{w}(\mathrm{t}) = \mathbf{C}\mathbf{x}(\mathrm{t}) \tag{2.34}$$

Here, **x**, **u**, **w**, and **θ** are the vectors of state, manipulated, output, and uncertainties, with sizes $n_x \times 1$, $n_u \times 1$, $n_w \times 1$, $n_\theta \times 1$, respectively. Accordingly, **A**, **B**, **C**, and **F** are the matrices of sizes $n_x \times n_x$, $n_u \times n_x$, $n_w \times n_x$, and $n_\theta \times n_x$. It is assumed that the matrices are time invariant and linear over the involved variables, and that **A** is non-singular (invertible). This discussion focuses on regulatory performances, therefore the input reference **r** is zero. For brevity, the manipulated variables in this section are referred to as the input variables. In most cases, all variables are scaled such that their absolute values are equal or less than 1.

For linear processes, the solution of control profile **u**(t) problem subject to (2.35)-(2.36) in terms of **x**(t) is unique (Anderson and Moore, 1990). The time domain representations, (2.16) and (2.17), are easily transformed back and forth to the complex domain **s** as follows:

$$\mathbf{sx(s)} = \mathbf{A(s)x(s)} + \mathbf{B(s)u(s)} + \mathbf{F(s)\theta(s)} \tag{2.35}$$

$$\mathbf{w(s)} = \mathbf{C(s)x(s)} \tag{2.36}$$

Furthermore, the transfer functions describing direct relationship between the manipulated variables **u**, the disturbances and uncertainties **θ**, and the outputs **w** can be derived as follows:

$$\mathbf{w(s)} = \mathbf{G(s)u(s)} + \mathbf{G_\theta(s)\theta s)\theta} \tag{2.37}$$

$$\mathbf{G(s)} = \mathbf{C(s)}\left(\mathbf{sI} - \mathbf{A(s)}\right)^{-1}\mathbf{B(s)} \tag{2.38}$$

$$\mathbf{G_\theta(s)} = \mathbf{C(s)}\left(\mathbf{sI} - \mathbf{A(s)}\right)^{-1}\mathbf{F(s)} \tag{2.39}$$

This transformation is also applicable to the frequency domain, where the notation **s** in (2.37)-(2.39) is replaced by **jω**. These expressions can be solved easily with linear algebra, in comparison to (2.33)-(2.34), which requires differential solution. Therefore, it is a standard procedure to transform the time domain state space to the frequency or complex domain (2.35)-(2.36), solve (2.37)-(2.39), then transform the solution back to time domain to extract the time profile **x**(t) and **w**(t). Based on these properties, the linear controllability analyses are often based on frequency domain specifications, such

as bandwidth ω_B (Anderson and Moore, 1990). Despite the attractive computation, it should be noted that the linear properties are only valid locally, mostly without consideration of process constraints.

The linear controllability analyses can be categorized into process structure and process dynamic analysis. The structural analyses are based on steady state calculation and focused on selecting the suitable pairing between input and output variables. The approaches along this line include the Relative Gain Array (RGA) and Niederlinski Index (NI), which are reviewed in section 2.7.1.1

The dynamic analyses investigate the quality of process dynamics, and at later development adopt the structural consideration. These analyses are based on the availability of the perfect control. Specifically for the disturbance rejection in (2.37), the perfect control requires $\mathbf{u(s)=G^{-1}(s)w(s)}$. In other words, the process inverse $\mathbf{G^{-1}(s)}$ should be available. The studies based on this concept believe that the process controllability is the inherent property of the process and should be independent of the controller.

Morari and Skogestad (1985) listed four fundamental limitations that prevent the perfect control, which are the Right Half Plane (RHP) zeros, the time delays, the constraints on manipulated variables, and the mismatch between the process and the model. A RHP zero in the process model would turn into the RHP pole in its inverse, which is equivalent to an unstable controller. A time delay would lead to an a-causal controller. The input constraints and model uncertainties prevent the accurate inversion of the process. The controllability assessments with the process-oriented approach have been concerned with the attainability of the perfect process control that is limited by these factors. Two approaches along this line are the Functional Controllability (Rosenbrock, 1970) and the Dynamic Resiliency (Morari and Skogestad, 1985), which are discussed further in section 2.7.1.2 and 2.7.1.3.

2.7.1.1 Linear Structural Controllability

The quality of process control depends on the compatibility of the manipulated (input) and the measured (output) variables. Based on this requirement, the Relative Gain Array (RGA) (Bristol, 1966) was introduced as a matrix of interaction measures for all possible SISO control loops in a process. It therefore indicates the interaction and preferable pairings between the manipulated variables and the output variables, which determines the control loops in a multi-loop SISO process control. It was originally defined at the steady state, or zero frequency condition ($\boldsymbol{\omega} = 0$ or $\mathbf{s} = 0$) (Bristol, 1966); but has been generalised since to frequency-dependent cases (Hovd and Skogestad, 1994; McAvoy, 1983; Stanley *et al.*, 1985; Wolff *et al.*, 1992b). The general formulation of RGA is therefore as follows:

$$\mathbf{\Lambda(G(s))} = \mathbf{G(s)} \times \left(\mathbf{G(s)}^{-1}\right)^{\mathbf{T}} \tag{2.40}$$

Here, $\times$ denotes the element-by-element multiplication. $\mathbf{G(s)}$ is a matrix transfer function of square process at a certain frequency, which is assumed stable and strictly proper. More importantly, the steady state gain $\mathbf{G(0)}$ is assumed nonsingular. The $\mathbf{ij^{th}}$ element of $\mathbf{\Lambda}$ is the ratio of the open-loop gain from input **j** to output **i** when all other loops are open; to the gain from input j to output i when all other loops are perfectly controlled. The relative gains $\boldsymbol{\lambda_{ij}}$ are the indicator of the sensitivity of the output **i** to the input **j**, where the value $\boldsymbol{\lambda_{ij}} \rightarrow 1$ with $0 \leq \boldsymbol{\lambda_{ij}} \leq 1$ is preferred.

Despite its considerable popularity, RGA does not necessarily provide a reliable variable pairing for processes larger than 2×2. In particular, this occurs when the feasibility of a variable pairing is entirely dependent on other control loops (Haggblom, 1997). To deal with this problem, a number of measurements that complement the RGA have been proposed, among them are the Block Relative Gain (BRG) (Manousiouthakis *et al.*, 1986) and the Niederlinski Index (Niederlinski, 1971).

The BRG is a feasibility measurement of block-decentralised control, where multiple input multiple–output (MIMO) control is considered for part of the process. The steady state BRG for MIMO control of subsystem $\mathbf{G_m(s)}$ is defined as follows:

$$\mathbf{\Lambda_m^B(G(s)) = G_m(s) \times \overline{G}_m(s)^{-1}} \tag{2.41}$$

Here, $\mathbf{\overline{G}_m(s)}$ is the effective gain matrix of subsystem $\mathbf{G_m(s)}$ when the rest of the process is perfectly controlled (Grosdidier and Morari, 1986).

Let $\mathbf{\widetilde{G}(s)}$ denotes the matrix obtained by setting to zero all elements of **G(s)** that do not correspond to an input-output pairing in a given (block) decentralized control structure. The Niederlinski Index (NI) as the frequency dependent measurement for the corresponding structure is defined as follows (Chiu and Arkun, 1990; Haggblom, 1997):

$$\mathbf{NI_{\widetilde{G}}(G(s)) = \det(G(s))/\det(\widetilde{G}(s))} \tag{2.42}$$

For a fully decentralised control structure, $\det(\mathbf{\widetilde{G}(s)})$ is equal to the product of the gains corresponding to the input-output pairings. Therefore, if the variable pairings are along the diagonal of **G(s)**, then $\det(\mathbf{\widetilde{G}(s)}) = \Pi_i g_{ii}(s)$. In that case, the subscript $\mathbf{\widetilde{G}}$ can be dropped, hence the Niederlinski Index becomes **NI(G(s))**(Haggblom, 1997).

These indices only recommend the suitable pairing between input and output variables, without further consideration of the corresponding process control dynamics. In the next two sections, the process control dynamics are assessed based on the process model **G(s)**.

2.7.1.2 Functional Controllability

The functional controllability (Rosenbrock, 1970) was defined based on the linear state space system in (2.35)-(2.36). The definition is that a linear system is functionally controllable if, given the smooth and causal output functions and zero initial conditions of the state variables, there exists a smooth manipulation profile that

generates the output functions. This definition also applies to both servo and regulatory control cases with little manipulation (Narraway *et al.*, 1991). For a square process, the functional controllability is achieved only if the process is invertible (Narraway *et al.*, 1991; Perkins and Wong, 1985). Meanwhile, non-square process requires that the rank of the transfer function matrix is equal to the number of inputs (Hovd and Skogestad, 1994). These relate to the existence of a perfect controller. Accordingly, the functional controllability analysis measures the extent of the physical limitations that prevents the implementation of the perfect controller.

The limitations include the presence of Right Half Plane (RHP) transmission zeroes, time delays, constraints on manipulated variables and process model mismatch. The RHP zeros correspond to unstable process inverse, hence limits the closed loop performance (Rosenbrock, 1970). Perkins and Wong (1985) characterised the effects of time delays with a parameter Δ_{min}, which is the minimum time before a trajectory for any $\mathbf{w}(t)$ can be specified. This measurement is calculated based on the magnitudes and locations of delays in the process transfer function.

The effects of the constraints in the manipulated variables can be characterised using the Singular Value Decomposition (SVD) of the transfer function $\mathbf{G(j\omega)}$. This transform the relationship between $\mathbf{u(j\omega)}$ and $\mathbf{w(j\omega)}$ into the following expression (Narraway *et al.*, 1991):

$$\mathbf{w}(\mathbf{j\omega}) = \mathbf{O}_r \mathbf{\Sigma} \mathbf{O}_l \mathbf{e}_u \tag{2.43}$$

where $\mathbf{O}_r$ and $\mathbf{O}_l$ are unitary orthonormal matrices made up of the right and left singular vectors of $\mathbf{G}(\mathbf{j\omega})$, $\mathbf{\Sigma}$ is a diagonal matrix of the singular values of $\mathbf{G}(\mathbf{j\omega})$ and $\mathbf{e}_u$ is a unit vector of the manipulated variable bounds.

In this analysis, the bounds of the manipulated variables are scaled to unity. Accordingly, the norm of the output responses are bounded by the maximum and minimum singular value of $\mathbf{G}(j\omega)$, σ_{max} and σ_{min}, as follows:

$$\sigma_{min}\|\mathbf{u}\|_2 \le \|\mathbf{w}(j\omega)\|_2 \le \sigma_{max}\|\mathbf{u}\|_2 \tag{2.44}$$

Therefore, the larger the singular values, the larger the region for output trajectories can be specified.

The ratio between the maximum and the minimum singular values is called the process condition number γ, with the definition given below:

$$\gamma \equiv \frac{\sigma_{max}}{\sigma_{min}} \tag{2.45}$$

This condition number represents the effects of the manipulated variable constraints. If γ is large, the outputs show a strong directionality based on the inputs. This deforms the shape of the output region and limits the specification of the trajectories as follows:

$$\frac{\|\partial\mathbf{u}\|}{\|\mathbf{u}\|} \le \gamma\frac{\|\partial\mathbf{G}\|}{\|\mathbf{G}\|} \tag{2.46}$$

Therefore, when the constraints on manipulated variables exist, the good control would require **large σ_{min}** and **small γ** (Perkins and Wong, 1985)**.**

Therefore, the consideration of functional controllability gives three separate indices Δ_{min}, σ_{min}, and γ; which characterise the attainability of perfect control. However, these are the 'yes-no'-type indicators. There is no obvious relationship between these indices, and no guidance to improve one while penalising the others. This is particularly the case for the condition number and singular value, since the measures are 'small' and 'large'. When the process is not functionally controllable, there is no indication of how far it is from the achievable performance. In the following section, the Dynamic Resilience concept, especially on its ability to provide a quality measure of the achievable performance, is reviewed.

2.7.1.3 Dynamic Resiliency

Dynamic Resiliency (Garcia and Morari, 1982) was defined as the quality of the regulatory and servo performance which can be obtained by feedback control. This concept is closely related to functional controllability, but the Dynamic Resilience expresses the process inherent limitation to attain the perfect controller. The index is based on the linear transfer function (2.37) and computed within the Internal Model Control (IMC) scheme. Due to the linearity assumption, the process properties in frequency domain, such as bandwidth ω, are prominent in this scheme. Most of the controllability measurements in this section refer to normalised variables with respect to their expected ranges (Wolff *et al.*, 1992b).

To generate the Dynamic Resilience index, the process transfer function **G(s)** is decomposed into an invertible matrix $\mathbf{G}^-$ and non-invertible matrix $\mathbf{G}^+$. The non-invertible matrix $\mathbf{G}^+$ contains all of non-minimum phase elements, such as time delays and right half plane zeros, which prevents realisable model inversions or functional controllability. The controller transfer function is given by $\mathbf{[G^-]^{-1}}$, $\mathbf{G^+(0)=I.}$ The limiting performances addressed within the Dynamic Resilience scheme are the time delays and the RHP zeros. The constraints on the manipulated variables, and the process-model mismatch are addressed in the similar way as with the functional controllability approach.

Holt and Morari (1985) presented various controllability bounds for processes that are not functionally controllable. The optimum noninvertible matrix $\mathbf{G}^+$ is the one that minimises the Integral Squared Error (ISE) and the Integral Absolute Error (IAE) of the output profiles as follows:

$$\mathrm{ISE} = \int_0^\infty \left|\mathbf{w}(t) - \mathbf{r}(t)\right|^2 dt \tag{2.48}$$

$$\mathrm{IAE} = \int_0^\infty \left|\mathbf{w}(t) - \mathbf{r}(t)\right| dt \tag{2.49}$$

Since the interest is in regulatory processes, then $\mathbf{r}(t) = 0$.

For a single-input single-output (SISO) closed-loop process with time delay τ, the upper bound of allowable bandwidth is $\omega_B<1/\tau$. The bandwidth of the process with RHP zeros z_{RHP} is $\omega_B<z_{RHP}$. Since fast dynamic responses corresponds to large bandwidth, the RHP zeros close to origin are generally avoided. If both RHP zeros and poles exist, it is important to ensure that the poles are far from the origin. The imperfect control of variable i subject to disturbance j is indicated by the magnitude $|[\mathbf{G}^{-1}(j\omega)\mathbf{G}_\theta(j\omega)]_{ij}|>1$.

The singular value analysis in this approach is performed on the process and disturbance transfer functions, $\mathbf{G}$ and $\mathbf{G}_\theta$. This determines the combination of manipulated variables and disturbances that cause the largest output variation. Morari (1983) introduced the corresponding minimum singular values $\sigma_{min}(G)$ and input magnitudes to judge the attainable process performance. A small singular value implies the requirement of large input magnitude, and therefore is undesirable.

Since $\sigma_{min}(G)$ is applicable only to unstructured uncertainty assumption, a structured singular value μ is proposed (Morari and Skogestad, 1985). This index is defined as the smallest perturbation that makes the process singular at each frequency. A large μ indicates high sensitivity, and small capacity to handle the structured perturbations while keeping the closed loop system feasible.

The direction of a disturbance to the process is indicated by the disturbance condition number defined as follows:

$$\gamma_\theta(G)=\frac{\left\|G^{-1}g_\theta\right\|_2}{\left\|g_\theta\right\|_2}\bar{\sigma}(G) \tag{2.50}$$

Here, g_θ is the element of G_θ. The disturbance condition number of G, $\gamma_\theta(G)$, indicates the amount of manipulation required to reject a unit disturbance, relative to the worst effect of the disturbance to the plant.

Skogestad and Morari (1987) subsequently introduced the Relative Disturbance Gain (RDG) to determine whether decentralised or full multivariable control can be used for disturbance rejection. The RDG is formulated based on the RGA (2.40), where **G** is replaced with G_θ to analyse the effects of the disturbances on steady state process. In this modification, a large relative gain element implies the extreme sensitivity to the disturbance, and is therefore undesirable. Hovd and Skogestad (1994) have extended the use of relative gain in dynamic processes, and introduced the Performance RGA (PRGA) to indicate any existing one-way coupling within the process as follows:

$$\Gamma(s) = G_{diag}(s) \times G(s)^{-1} \tag{2.51}$$

Here, G_{diag} is a matrix containing only the diagonal element of **G**.

Subsequently, the Closed-Loop Disturbance Gain (CLDG) is defined as follows:

$$\beta(s) = \left(GG^{-1}G_\theta\right)G_\theta^{-1} \tag{2.52}$$

Zhao and Skogestad (1997) later introduced the Partial Disturbance Gain (PDG) to measure the disturbance gain for a process under partial control as follows:

$$\beta_p(s) = \left[\left[G^{-1}\right]_{11}\right]^{-1} G^{-1}G_\theta \tag{2.53}$$

Trierweiler and Engell (1997) proposed the Robust Performance Number (RPN) and the Robust Performance Number of a Process Set (RPPN). The number indicates how potentially difficult it is for a given system to achieve the desired performance robustly. They both reflect the attainable performance of a process and its degree of directionality. RPN is used for nominal systems, whereas the RPPN is used for nonlinear and uncertain process with the scaled model.

Several different linearised models, depending on the selection of linearisation points, may represent one process. This raises the notion of the gap metric (Galan *et al.*, 2003; Samyudia *et al.*, 2000; Samyudia and Lee, 2002 and the references therein), which measures the closeness of two linear open loop models. The gap is defined as follows:

$$\delta(\mathbf{G_1}, \mathbf{G_2}) = \max\left\{\vec{\delta}(\mathbf{G_1}, \mathbf{G_2}), \vec{\delta}(\mathbf{G_2}, \mathbf{G_1})\right\} \tag{2.54}$$

$$\vec{\delta}(\mathbf{G_1}, \mathbf{G_2}) = \inf_{\mathbf{Q} \in H_\infty} \left\| \begin{bmatrix} \mathbf{M}_1 & \mathbf{N}_1 \end{bmatrix} \mathbf{Q} \begin{bmatrix} \mathbf{M}_2 & \mathbf{N}_2 \end{bmatrix} \right\|_\infty \tag{2.55}$$

Here, δ is the gap metric, while G_1 and G_2 are the linear open loop models of interest. The directed gap $\vec{\delta}(G_1, G_2)$, is the function of the left and right coprime factors of each models, M and N respectively; where **Q** is a finite weighting matrix. The gap value close to 0 indicates that the models are close, while the value closer to 1 indicates otherwise. The motive behind this metric is to quantify the perturbations of an open loop system that still maintains closed loop stability. In other words, if two open loop systems are close, they will show similar behavior in the closed loop.

These controllability measurements have been applied to various cases. Skogestad and Wolff (1992) presented the evaluation of process sensitivity to sinusoidal disturbance for open-loop, decentralised control, partial control and regulatory control. Wolff *et al.* (1992a; 1992b) demonstrated the controllability assessments at a FCC reactor and a Hydrodealkylisation of toluene (HDA) process. Weitz and Lewin (1996) introduced the Disturbance Cost (DC), which is similar in principle to the disturbance condition number. Zhao and Skogestad (1997) applied the PGA to select the best control configuration of continuous bioreactors. Hovd and Skogestad (1994) investigated the pairing of manipulated variables and the output variables using RGA. Hernandez and Jimenez (1999) analysed the controllability properties of thermally coupled distillation sequences. Cao *et al.* (1997) simultaneously assessed various control schemes on a CSTR and a HDA process. The gap metric has been applied to multi-linear model based control design and control structure analysis of process recycle streams (Galan *et al.*, 2003; Samyudia *et al.*, 2000; Samyudia and Lee, 2002).

There are several limitations of these approaches. Most of the indices are based on the frequency domain specification. Meanwhile, in practice, time domain performances are more favorable. Although linearity leads to elegant analytical expressions, it only applies to a narrow operating region. In addition, it assumes a square transfer function matrix **G** related to the same number of inputs and outputs, which is not a realistic condition. The usage of the process and disturbance transfer functions **G** and $\mathbf{G_\theta}$ only provide the open-loop performance measurements, which is biased towards sequential process control design scheme. The disturbance rejection measurements assume that the uncertainties are bounded and unstructured. With this assumption, the effect of disturbance is measured as its worst-case effect, which leads to conservative specifications.

Further to these problems is the difficulty of interpreting the results of these controllability analysis techniques. For instance, identifying what is too large a value of a condition number is not clear, and therefore, can be misleading. The indices typically consider one of the performance limitations in isolation, which makes it difficult to understand the combined effect of these limitations on the process controllability. There is no guidance on how to balance conflicting specifications, or how to compare several processes with different structures.

Therefore, further works are still required to produce indices that are easy to interpret and applicable to non-square process. These indices shall also deal directly with time domain specifications and address the combined effects of as many fundamental limitations as possible. Therefore, a controllability analysis that is directly applicable to the nonlinear model of the plant would be more useful. This approach is reviewed in the following section.

2.7.2 Nonlinear Controllability Analysis

The existing nonlinear controllability analyses can be classified into the analytical and the optimisation methods. In comparison to the amount of effort that has been placed in the assessment of linear controllability, there have been few publications about

nonlinear controllability. Some insights to the analytical nonlinear controllability analysis are given below, followed by the review of the optimisation methods.

2.7.2.1 Analytical Methods

Despite the limitation of 'yes-no' type diagnosis, the functional controllability concept has been extended to nonlinear cases. Hirschorn (1979) and Singh (1982) derived the sufficient conditions for the functional controllability of nonlinear processes. Subsequently, the necessary and sufficient conditions for functional controllability are presented for SISO (Tsinias and Kalouptsidis, 1983) and MIMO (Li and Feng, 1987) cases.

Daoutidis and Kravaris (1991) introduced the unstable zero dynamics and the relative order analogous to the RHP zeros and dead time in a nonlinear process. The unforced zero dynamics of a nonlinear process is given by the dynamics of the minimal order realisation of the process inverse. This inverse is determined by the desired process output trajectory and the number of relative orders minus one.

The relative degree is the lowest order derivative of the output subject to the available input variables. Small relative order is preferable, since it indicates more direct effect of the input to the output, more aggressive initial response, and smaller the dead time. Whenever possible, the derivative matrix is rearranged so that the minimum relative order for each row lies on the diagonal to give the minimum input-output pairing, while the interactions with other alternative pairings are analysed through the off-diagonal elements

Similarly to the usage of the Relative Gain Array (RGA) in linear processes, the static RGA is also used for nonlinear processes (Mijares *et al.*, 1985). Manousiouthakis and Nikolau (1989) extended the Block Relative Gain (BRG) concept into the steady state Nonlinear BRG (NBRG) and the Dynamic NBRG (DNBRG) for 2 x 2 process, as given in (2.56).

$$\begin{aligned} DNBRG_l &= G_{11}(G^{-1})_{11} \\ DNBRG_r &= (G^{-1})_{11}G_{11} \\ \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= G(u_1, u_2) \end{aligned} \tag{2.56}$$

The element of DNBRG = 1 suggests that the respective pair of manipulated and output variables are independent from other loop, while any value > 1 indicates an interaction. It is also shown that the steady state NBRG is the minimum condition number for nonlinear systems (Manousiouthakis and Nikolau, 1989).

The selection of the manipulated and measured variable pairings to guarantee stability has also been addressed through the formal connection between the process thermodynamics and the passivity theory of nonlinear control. A system is rendered passive when the rate of change of the generalised energy stored in the system (storage function) is bounded from above by the product of the manipulated inputs and the measured outputs. This is known as the supply rate inequality for passive system. The specific property is that the feedback interconnection of a passive system and a strictly passive system is asymptotically stable. Ydstie and coworkers (Farschman *et al.*, 1998; Hangos *et al.*, 1999; Ydstie and Alonso, 1997) employed this concept to select the pairing such that the interconnection between the process and the controller is passive. A detailed understanding of process thermodynamics is required to derive the analytic supply rate inequality. So far, the resulting 'yes-no'-type indicator applies locally around a stationary, minimum entropy point.

Except on the passivity approach, the general guideline on whether an element should be big or small for these nonlinear controllability analyses has been provided. However, it is not obvious what is considered too big or too small a value of any element in these analyses. Similar to the linear versions, these indices require experience for interpretation. Also, there is no guidance on how to produce a better design. None of these analytical techniques addresses more than one of fundamental limitations on controllability at a time.

2.7.2.2 Optimisation Methods

The optimisation methods perhaps are the most successful methods to integrate the process design and controllability. The important factor is their ability to consider multiple specifications, either as constraints or objectives within the formulation of the optimisation problem (2.27). This capability enables the method to quantify process controllability, or to integrate one or more controllability indices into process control synthesis, preferably to satisfy the process economy. These have been made possible by the advances in computational hardware and optimisation tools (Walsh and Perkins, 1996). This section reviews the quantification of process controllability using optimisation methods that lead to introduction of integration between process control and process design.

Dimitriadis and Pistikopoulos (1995) extended the steady state flexibility test and flexibility index problem (Swaney and Grossman, 1985a) to a fixed structure dynamic process. In this extension, the process dynamics are represented by a set of DAE, subject to a single time-varying uncertainty. Accordingly, the dynamic feasibility problem of the process is as follows:

$$\chi(\mathbf{z},\mathbf{y}) = \max_{\theta\in\Theta} \psi(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t),t) \tag{2.57}$$

$$\mathbf{s.t.} \quad \psi(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t),t) = \min_{z\in Z} \upsilon$$

$$\mathbf{s.t.} \quad \mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t),t) = \mathbf{0} \qquad \mathbf{i}\in\mathbf{E}$$

$$\mathbf{g_j}(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t),t) \le \upsilon \qquad \mathbf{j}\in\mathbf{I}$$

$$\boldsymbol{\Theta}(t) = \{\boldsymbol{\theta}(t) \mid \boldsymbol{\theta}^{\mathbf{L}}(t) \le \boldsymbol{\theta}(t) \le \boldsymbol{\theta}^{\mathbf{U}}(t)\}$$

In this formulation, $\mathbf{x}$ and $\dot{\mathbf{x}}$ are the vectors of state variables and their derivatives, $\mathbf{u}$ is the vector of manipulated variables, and $\boldsymbol{\theta}(t)$ is the uncertainty profile. If $\chi \le 0$, then the proposed design is dynamically feasible in $\boldsymbol{\Theta}(t)$. Otherwise, the process is not feasible. The formulation establishes that for every possible profile of the uncertain vector $\boldsymbol{\theta}(t)$, there is at least one profile of the manipulated variable $\mathbf{u}(t)$, such that the process feasibility constraints are satisfied within the time horizon $t \in [0, t_f]$.

The dynamic feasibility index is formulated as the following optimisation problem:

$$DF(\mathbf{z},\mathbf{y}) = \max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \delta \tag{2.58}$$

$$\textbf{s.t} \quad \chi(\mathbf{z},\mathbf{y}) = \max_{\boldsymbol{\theta}(t)\in\Theta(\delta,t)} \min_{u(t)\in U(t)} \max_{j\in I, t\in[0,H]} g_j(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t)) \le 0$$

$$\textbf{s.t..} \quad \mathbf{h_i}(\mathbf{z},\mathbf{y},\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\mathbf{w},\boldsymbol{\theta}(t),t) = \mathbf{0} \qquad \mathbf{i}\in\mathbf{E}$$

$$\delta \ge 0 \qquad \mathbf{j}\in\mathbf{I}$$

$$\boldsymbol{\Theta}(t) = \{\boldsymbol{\theta}(t) \mid \boldsymbol{\theta}^{\mathbf{L}}(t) \le \boldsymbol{\theta}(t) \le \boldsymbol{\theta}^{\mathbf{U}}(t)\}$$

Qualitatively, the Dynamic Flexibility Index (DF) represents the largest scaled deviation of the uncertain parameter profile that the design can tolerate while remaining feasible within the time horizon $t\in[0, t_f]$.

One main challenge of the optimisation based controllability analysis, including dynamic flexibility test and flexibility index problems, is solving the process dynamics within the optimisation problem. There are two known strategies for this purpose, namely the sequential and simultaneous methods.

The sequential methods parameterise the decision variables and the manipulated variables. The process dynamics are solved using a DAE solver for the given values of decision variables. The DAE solutions are then used to evaluate the objective and the remaining constraints at each optimisation step.

When handling optimal control problems, the sequential strategy may use the Control Vector Iteration (CVI) or Control Vector Parameterisation (CVP) technique. CVI requires solution of the Euler-Lagrange equations and minimisation of the Hamiltonian, while the CVP involves the repeated DAE solution. The computational costs of both methods are expensive, but the numerical error is relatively small.

In the simultaneous method, the differential equations are converted into algebraic residual equations using discretisation methods such as orthogonal collocation (Villadsen and Michelsen, 1978). This technique approximates the profiles of the state and manipulated variables at a number of collocation points that are determined based on the zeros of the corresponding Lagrange polynomials. In addition, orthogonal

collocation on finite elements method divides the time horizon into a pre-specified number of finite elements and uses a separate approximation for each element. The continuity of the state and manipulated input profiles is enforced at the beginning of each element.

This procedure transforms the dynamic optimisation problem into a large-size Nonlinear Programming (NLP) problem (Cuthrell and Biegler, 1987). When there are n_i number of DAE, n_x number of state variables, n_u number of manipulated variables, n_{el} number of element with n_{col} number of collocation point, then the original DAE is converted into $n_i \times n_e \times (n_{col}+1)$ equations. Meanwhile, the optimisation problem has the additional $(n_x+n_u) \times n_e \times (n_{col}+1)$ decision variables. The element lengths and the allowed breakpoints for control profiles discontinuities can be included as the decision variables (Logsdon and Biegler, 1989). In order to efficiently solve the large-scale NLP problem, Tjoa and Biegler (1991) proposed the Sequential Quadratic Programming (SQP) method. Cervantes and Biegler (1997) proposed further size reduction of the SQP problem by exploiting the overall sparsity of the collocation matrix.

Process syntheses are likely to involve decisions about the existence of particular units or streams, the connection or disconnection of particular units, and the cause and effect relationship among the parameters. This problem generates a set of possible process structures (superstructure) that may satisfy the design objectives, such as disturbance rejection. To select the best structure, Georgiou and Floudas (1989) proposed a generic rank of process structural matrix as an index of structural controllability. In this approach, the process dynamic model is transformed into an augmented structural matrix. This matrix represents the relationships between variables in a particular process structure, with no dependency to the operating conditions and system non-linearity. The generic rank of the matrix is computed through the solution of an Integer Linear Programming (ILP) problem. The full rank indicates the feasibility of the structure to satisfy a specified objective, and weeds out the unfeasible structures or excess measurement variables. Similar approach is reported by Hopkins *et al.* (1998),

where the cause and effect matrix (CEM) method is used to determine the respective Output Structural Controllability (OSC).

Chenery and Walsh (1998) determined the best achievable control performance of a process over a set of admissible controllers subject to a given set of disturbances. The controllability index is an L_1-norm of the objective function, which is minimised with the selected static controllers along with their operating points.

Zheng and Mahajanam (1999) proposed a quantitative controllability index, which is principally the minimum over design required to guarantee process feasibility. The main idea is that the dynamics due to disturbances has less effect if the process has a bigger capacity to absorb them. Therefore, the system controllability for all the expected disturbances can be determined by examining the installation costs of the additional surge tanks into the process. The minimum cost related to the volume of the surge tanks (**v***) becomes the proposed controllability index, which is solved subject to the allowable dynamics of control variables, manipulated variables and disturbances, as well as process characteristics and control structures. The main property of **v*** is that a process is controllable if and only if $\mathbf{v}^* \geq 0$. Consequently, **v** is bounded only if the steady state control problem is feasible, or if the closed-loop problem is asymptotically stable. Since **v** can be represented by its installation cost, the index allows the comparison between several alternatives of process control design in economic terms.

A new alternative for process controllability assessment is proposed by Vinson and Georgakis (1998; 2000). They introduced the Output Controllability Index (OCI) as a steady state, input-output controllability measurement of a process. It quantifies the ability of the process to reach the full range of desired output values in the presence of expected process disturbances, within the limited range of its available inputs. A geometric software (Veres *et al.*, 1995) is used to characterise inputs, outputs and disturbances dynamics as multi-dimensional polyhedrons, and calculate their multi-dimensional volumes to determine the index. Veres *et al.* (1995) developed the software to characterise model uncertainties as geometric shapes, such as polytopes and ellipsoids, then transformed these shapes through a linear model, to assess feasible

models and to track changes of linear system dynamics. Subsequently, Georgakis *et al.* (2001) embedded the OCI into the optimisation of process control, to assess the inherent dynamic operability. Meanwhile, Ierapetritou (2001) performed similar geometric calculations for the Feasibility Index (Grossmann *et al.*, 1983) of steady state and convex problems.

Hovd and Braatz (2000) proposed the computation of disturbance rejection indices for non square processes. The problem is formulated as a non-convex optimisation problem, which requires the use of global solution for large problems. To reduce the problem size, Kookos and Perkins (2003) reformulated the problem into the mixed-integer linear programming formulations and extended the measures to nonlinear problems.

These controllability measures show that the optimisation methods provide the flexible formulations and easily interpreted results, in comparison to the non-optimisation based counterparts. Since the analyses are performed in the time domain, the performance specifications and the limitations on the control performance, such as the constraints on state, manipulated and output variables, are easily incorporated into the formulation. Therefore, the feature can be employed further to produce the controllability assessment that is applicable for a given set of alternative process and controller structures. The feature also accommodates the integrated considerations of process economy and controllability in the early stage of process design and synthesis. It should be noted that these involve high computational costs due to the DAE solution. While the advances of computational technology would relieve some of the loads of the problem, the efficient solutions still depend on developed strategies within the optimisation framework. These integrated frameworks are reviewed in the next section.

2.8 Integrated Economic and Controllability Considerations in Process Design and Synthesis

2.8.1 Existing Studies

The simultaneous consideration of process control design and process economy using multi objective frameworks can be traced back to Lenhoff and Morari (1982). Palazoglu and Arkun (1986; 1987) also proposed the process design with robustness indices describing the dynamic operability of fixed structure flow sheets. Luyben and Floudas (1994a; 1994b) used a multi objective approach to simultaneously consider the open-loop controllability and process economics of process synthesis. Papalexandri and Pistikopoulos (1994a; 1994b) applied the generic rank approach (Georgiou and Floudas, 1989) to a synthesis/retrofit case of heat exchanger network (HEN). The problem was formulated as a MINLP problem to minimise the total annualised synthesis/retrofit cost subject to process constraint and disturbance rejection specifications.

These studies address the process dynamics indirectly. The approaches include the analysis of transfer function matrices (Lenhoff and Morari, 1982; Luyben and Floudas, 1994a; b; Palazoglu and Arkun, 1986; 1987) and the representation of the dynamic response as function of design variables (Papalexandri and Pistikopoulos, 1994a; b).

The integrated assessments of process economy and dynamic performances without the optimisation framework were reported by Luyben and coworkers (Elliot and Luyben, 1995; 1997; Lyman *et al.*, 1996). They designed several alternative process control structures separately based on economic objective. Subsequently, the respective dynamic performances were assessed and ranked based on the ISE (Lyman *et al.*, 1996), and the frequency domain specifications such as bandwidth, magnitude ratio, phase angle and peak log modulus (Elliot and Luyben, 1995; 1997).

Perkins *et al.* (1989) were among the first to directly assess the effects of process dynamics on the process economic performances within an optimisation framework. Their method was based on the optimal operation of a chemical process, and the necessary back off to avoid the constraint violations due to the effects of disturbances and uncertainties. The method used a linearised state space process model, which permits the generation of first order estimates of relevant quantities (Gannavarapu, 1991). The variations of each variable were used to estimate the required back off for ensuring the feasibility, as well as to estimate the change of the process economics. The economic analysis was carried out at the expected disturbance frequencies and amplitudes. In this work, they assumed a perfect control to avoid analysing the effect of controller to the process dynamics.

A simple way to consider the economic effects of disturbances is to look at their impacts on the variations in process variables. This is indicated by the slack variable of the active constraints. If under nominal operating conditions the slack variable value of an active constraint is non-zero, then the constraint is violated. To remove the violation, the operating point should be moved by a sufficient distance into the operating region. This is the back off move, which is represented by the Lagrange multiplier λ_i, associated with the slack variable δ_i. Narraway *et al.* (1991) solved the optimum linear steady state problem, then estimated the associated economic penalty as follows:

$$\Phi = \Phi_0 - \sum_{i=1}^{n_i} \lambda_i \max \delta_i \tag{2.59}$$

Here, Φ is the back off objective value and Φ_0 is the optimum objective value at the steady state optimum. The difference between the objectives is determined over a number of active constraints n_i, through the Lagrange multiplier and the slack variable of the i^{th} active constraint, λ_i and δ_i respectively.

The penalty estimation is performed by transforming the linear process model, constraints, and the slack variables computation into the frequency domain. This

accordingly transforms the optimisation problem into a matrix algebra, which yields the maximum values of the slack variables as the function of the disturbances $\delta(\theta)$. In the subsequent research, Narraway and Perkins (1994) extended the approach to nonlinear process and included the controller structure selection. Since the matrix computation was not applicable anymore, the problem was formulated as MINLP. The corresponding integer variables indicated the pairing between the manipulated and controlled variables, and the problem was solved using the OA/ER/AP method (Viswanathan and Grossmann, 1990). The difficulty in finding the optimum solution was reported, which was due to nonlinearity and nonconvexity prominent in process dynamic solutions.

Bahri (1996) developed the Dynamic Operability Framework for operability assessment and process synthesis based on the back off approach for both linear and nonlinear dynamic processes. The framework is the continuation from previous work by Bandoni *et al.* (1994) and Bahri *et al.* (1996a) on the steady state flexibility analysis of a chemical process. The objective is to maximise the process economy subject to the feasible regulatory dynamics. Therefore, the economic penalty is determined by the distance between the steady state optimum and the dynamic operating point, which are calculated based on nonlinear steady state and nonlinear dynamic models, respectively. The associated multi-objective problem is reduced to a single-objective one by assigning the dynamic flexibility set as constraints. The affecting disturbances and uncertainties are characterised as a set of step functions with uniformly distributed magnitudes, and the worst-case disturbance combination causing the constraint violations is calculated automatically. The corresponding dynamic optimisation method is solved using the orthogonal collocation on finite elements method (Logsdon and Biegler, 1989). The control schemes are the pre-designed multi-loop PI and the multivariable controllers, which are fine-tuned within the framework.

The process synthesis decides the existence of process streams that define the process configuration, and the selection between different controller schemes. The optimum structure is selected simultaneously with the optimum parameters, based on the

process economy. The corresponding MINLP problem is solved using the OA method. The relevant application in solving industrial process structure selection problem is reported in Ngo and Bahri (1996) and Bahri *et al.* (1997). Bahri *et al.* (1996b) included the Integral Squared Error (ISE) as a weighted constraint.

The developments include the introduction of the recovery factor, as well as the ratio between the amount of penalty recovered with control, to the penalty with no control Figueroa *et al.* (1996). (Figueroa, 2000) further applied the back off approach to a variable structure control case.

Still in the same corridor, Vu *et al.* (1997) applied the back off problem to switchability case of a fixed process structure. The switchability deals with problems involving changeover policy, startup and shut down operations. This case seeks for the best trajectories of the control variables and the profiles of the state variables by minimising the ISE of the process variables for a fixed time horizon. The dynamic optimisation problem is solved simultaneously using the orthogonal collocation in finite element approach (Logsdon and Biegler, 1989).

The advantage of the back off approach is that they determine the cost increase associated with moving to the back off position, which is due to the uncertainties and disturbances. A limitation of this approach is that it leads to conservative design since the framework considers the worst-case uncertainty scenario, although the probability of the worst-case uncertainty may not be high.

In a parallel study, Mohideen *et al.* (1996) applied the dynamic flexibility test problem (Dimitriadis and Pistikopoulos, 1995) to an optimal process control design under uncertainty. The flexibility aspects are assessed simultaneously with the control specifications to find the economic optimum that satisfies all of the constraints for a given set of uncertainties and disturbances on a closed-loop process. The flexibility is assigned as a set of constraints (Bahri *et al.*, 1996b). The control scheme is defined as multi-loop PI controller, and controller parameters are optimized within the framework. The corresponding stochastic and dynamic programming is solved using

the Parametric Programming and the orthogonal collocation methods, respectively. The MINLP associated with the process and controller structures is solved using the GBD method (Geoffrion, 1972).

Schweiger and Floudas (1998) generated a set of trade-off solutions between the economy and controllability (in terms of ISE) in process synthesis. This multi-objective problem is solved by assigning the ISE as a weighted point constraint to the optimisation problem, and solving the problem successively at various weighting values – a procedure known as the ε-constraint method. The disturbance affecting the process is a single sinusoidal function, and the corresponding process dynamics are solved sequentially with control parameterisation method. The controller schemes are multi-loop PI controllers that are tuned within the framework. The MINLP problem associated with the process synthesis is solved using the OA and GBD method.

2.8.2 Analysis of the Existing Studies

These studies, especially the works of Bahri (1996), Mohideen *et al.* (1996) and Schweiger and Floudas (1997), demonstrate the systematic and integrated assessments of process synthesis, design and control based on the optimisation frameworks. The principal characteristics of the problems are the solution of process dynamics and the respective controllability index. These are subject to a given characterisation of disturbances and uncertainties, as well as the existence of both continuous and discrete decisions associated with process synthesis. The dynamics constitutes the major computational cost within the optimisation. The multi-objective nature of the problems is handled by assigning the dynamic flexibility, or the controllability condition, as the constraints within the frameworks.

Within these frameworks, the ISE (Bahri *et al.*, 1996b; Schweiger and Floudas, 1997) is easier to interpret in comparison to the dynamic flexibility function (Bahri *et al.*, 1996b; Mohideen *et al.*, 1996). However, ISE only represents one profile at a time. For a multivariable process that has several measured and constrained output variables, it is not yet clear which variables should be assessed with ISE. This is an important

decision, because different process structures may activate different dynamic profiles and activate different constraints. If the variable assessed with ISE happens to be inactive at a certain structure, for instance, because the corresponding equipment is not exist, then the corresponding ISE result might be confused with the 'small ISE', which is misleading. Furthermore, ISE only quantifies the dynamic profile against a point reference, and does not by itself guarantees the process feasibility or controllability (Bahri, 1996).

On the other hand, the dynamic flexibility function guarantees flexibility and controllability, but does not quantify them. It should be used along with the complementary dynamic flexibility index, which identifies the critical disturbance combination, and with the calculation of the corresponding penalty. In Bahri (1996), the economic penalty is calculated relative to the optimum steady state objective, and associated with the maximum magnitude of the critical dynamic profile. Over the time horizon, this indicates that the dynamic is bounded. However, it does not assess the quality of the control action as ISE does. These conditions show that a controllability index that can both guarantee and quantify multivariable controllability is yet to be developed, and is worth further investigation.

The selection of process and controller structure gives rise to MINLP problem, which were reportedly solved using the OA (Bahri *et al.*, 1996b) and GBD (Mohideen *et al.*, 1996; Schweiger and Floudas, 1997) techniques. Narraway and Perkins (1994) reported the difficulty of finding an optimum solution using the OA method due to solutions of nonlinear and nonconvex dynamics, which may indicate its inherent unsuitability to solve the problem. This fact requires further investigation into the nature of the MINLP solutions, and the practical improvement to solve common dynamic problems.

Lastly, the disturbance characterisations in the reported studies so far are limited to the scalar (Narraway and Perkins, 1994; Narraway *et al.*, 1991), sinusoidal (Schweiger and Floudas, 1997) and step (Bahri, 1996) functions. The corresponding assessment might either fail to anticipate all the possible dynamics within the operating domain, or

otherwise be too conservative. Therefore, the consideration of the general characterisation of disturbances in the controllability assessment also requires further investigation.

In brief, the existing controllability assessments would benefit from the further development of a multivariable controllability index, dynamic MINLP solutions, and general characterisation of disturbances. In the next section, the problem formulation to address these aspects as the objective of this thesis is presented.

2.9 Extension of Dynamic Operability Framework

This study addresses the gaps reviewed above through the extension of the original Dynamic Operability Framework (Bahri, 1996) with the incorporation of a formal controllability measurement. This takes into account the multivariable dynamic performances, the solution of the MINLP problem for nonlinear and nonconvex cases, and the accommodation of general disturbance characterisations.

One potential candidate for the controllability index is the characterisation of the feasible operating region and dynamic profiles as a multi-dimensional space. This approach is motivated by the Output Controllability Index (Vinson and Georgakis, 1998; 2000) approach, which is originally developed to determine a steady state output controllability index of a process system. This approach captures the effects of inherent interactions of multivariable process to the relationship between the disturbances and the measured outputs, in contrast to separate quantification's such as ISE. In this thesis, its extension and application to nonlinear dynamic processes, and the implication of its use on different, but fixed process structures are investigated in Chapter 3 and 4.

This development is further extended to cover the process synthesis, specifically to select the optimum process and controller structures from a given possible set of structures (superstructure). This includes the investigation of the existing MINLP solutions at both steady and dynamic states. The focus is on practical improvement of

convergence and optimality, in the presence of nonlinearity and nonconvexity. This development is discussed at length in Chapter 5.

A potential approach to address the disturbance characterisation and the conservatism issues, is the control unfalsification concept (Kosut, 2001; Safonov, 1996; Safonov and Tsao, 1997). In this concept, the process designs are allowed to adapt according to the severity of disturbance effects over time. This is realised by performing the assessment within an optimisation window that progresses sequentially over time. Rather than assuming some form of disturbance characterisations, the disturbance is sampled within the window, and the corresponding process responses are assessed. The resulting design may be updated as the window progresses, depending on the effects of the disturbances. The incorporation of this feature within the Dynamic Operability Framework is presented in Chapter 6.

The framework development is accompanied by relevant academic case studies. The ultimate demonstration is given at chapter 7, which is an industrial case of systematic controllability assessment of an industrial five-effect liquor-burning evaporator within an Alumina refinery. This application highlights the contribution of the proposed controllability assessment framework in bridging the process design methodologies with the industrial implementation.

2.10 Conclusion

This chapter has presented an overview of the existing operability assessment, especially the controllability, in process control design and synthesis. It has been shown that the controllability considerations lead to process designs that are not only economically viable, but also effectively employ the available resources to maintain the design specifications despite the variations in the process. This important concern of engineering design can be accommodated systematically at the design stage with the use of optimisation algorithms. The numerical tools would assist the assessment of

process controllability in terms of parameters and structure, simultaneously with process economy, subject to the known affecting disturbances.

The proposed algorithm is the extension of the Dynamic Operability Framework (Bahri, 1996), which originally addresses the fixed flexibility and controllability of process design. The extension would cover the incorporation of a controllability index that captures the multivariable process dynamics, the efficient selection of optimum process and controller structure, and the general disturbance characterisations. These aims are pursued step-by-step in the next four chapters, and finally applied to an industrial case study.

2.11 Nomenclature

Acronyms		
BB	:	Branch and Bound
BRG	:	Block Relative Gain
CEM	:	Cause and Effect Matrix
CLDG	:	Closed-Loop Disturbance Gain
CSTR	:	Continuous Stirred Tank Reactor
CVI	:	Control Vector Iteration
CVP	:	Control Vector Parameterisation
DAE	:	Algebraic Differential Equations
DC	:	Disturbance Cost
DF	:	Dynamic Flexibility Index
DNBRG	:	Dynamic Nonlinear Block Relative Gain
DP	:	Discrete Programming
ESF	:	Expected Stochastic Flexibility
FCC	:	Fluid Catalytic Cracking
FI	:	Flexibility Index
GBD	:	Generalised Bender Decomposition
HDA	:	Hydrodealkylation
HEN	:	Heat Exchanger Network
HSS	:	Hammersley Sequence Sampling
IAE	:	Integral Absolute Error
ISE	:	Integral Squared Error
IP	:	Integer Programming
ILP	:	Integer Linear Programming
IMC	:	Internal Model Control
LHS	:	Latin Hypercube Sampling
MC	:	Monte Carlo
MILP	:	Mixed Integer Linear Programming
MINLP	:	Mixed Integer Nonlinear Programming
MIP	:	Mixed Integer Programming
MIMO	:	Multiple-Input Multiple–Output
NBRG	:	Nonlinear Block Relative Gain
NI	:	Niederlinski Index
NLP	:	Nonlinear Programming
OA	:	Outer Approximation
OA/ER	:	Outer Approximation with Equality Relaxation
OA/ER/ AP	:	Outer Approximation with Equality Relaxation and Augmented Penalty
OCI	:	Output Controllability Index
ODE	:	Ordinary Differential Equations

Acronyms		
OSC	:	Output Structural Controllability
PDF	:	Probability Density Functions
PGA	:	Partial Disturbance Gain
PRGA	:	Partial Relative Disturbance Gain
RDG	:	Relative Disturbance Gain
RDS	:	Reduced Dimensional Stochastic
RGA	:	Relative Gain Array
RHP	:	Right Half Plane
RI	:	Resilience Index
RPN	:	Robust Performance Number
RPPN	:	Robust Performance Number of a Process Set
SISO	:	Single-Input Single-Output
SF	:	Stochastic Flexibility
SQP	:	Sequential Quadratic Programming
SVD	:	Singular Value Decomposition

Variables		
$\mathbf{A, B, C, F}$	:	Linearised state-space matrices
$\mathbf{E}$	:	Set of equality constraints
$\mathbf{E}_{\theta}$	:	Expected value with respect of θ
$\mathbf{e}_u$	:	Unit vector of manipulated variable bounds.
$\mathbf{H}$	:	Constant, linearised state-space matrices
$\mathbf{h}_c$	:	Vector of constant matrices
$\mathbf{G}$	:	Process transfer function
$\widetilde{\mathbf{G}}$	:	Matrix containing all elements of **G** corresponding to an input-output pairing only
$\mathbf{G}^-$	:	Invertible component of **G**
$\mathbf{G}^+$	:	Noninvertible component of **G**
$\mathbf{G}_{11}$	:	Element row 1 column 1 of **G**
$\mathbf{G}_{diag}$	:	Diagonal matrix of **G**
$\mathbf{G}_m(s)$	:	Subsystem of **G**
$\overline{\mathbf{G}}_m$	:	Effective gain matrix of subsystem $\mathbf{G}_m$
$\mathbf{G}_{\theta}$	:	Disturbance transfer function
$\mathbf{g}_{\theta}$	:	Element of $\mathbf{G}_{\theta}$
$\mathbf{g}_j$	:	Inequality constraints
$\mathbf{H}_{\infty}$	:	H_∞ norm
$\mathbf{h}_i$	:	Equality constraints, process DAE model
$\mathbf{I}$	:	Set of inequality constraints
$\mathbf{I}_A^k, \mathbf{I}_j$	:	Active constraints
$j\omega$	:	Frequency domain
$\mathbf{\Lambda}_m^B$	:	Block Relative Gain

Variables		
M	:	Large scalar values assigned to activate or deactivate constraints
$\mathbf{M_1}$	:	Left coprime factor of process transfer function G_1
$\mathbf{M_2}$	:	Left coprime factor of process transfer function G_2
$\mathbf{N_1}$	:	Right coprime factor of process transfer function G_1
$\mathbf{N_2}$	:	Right coprime factor of process transfer function G_2
$NI_{\tilde{G}}(G)$	:	Niederlinski Index
n_i	:	Number of the active constraints
n_θ	:	Number of the uncertainty parameters
n_s	:	Number of possible states in the process
n_z	:	Number of the design parameters
O	:	Set of objective functions
$\mathbf{O_l}$	:	Unitary orthonormal matrix made up of the left singular vectors of **G**
$\mathbf{O_r}$	:	Unitary orthonormal matrix made up of the right singular vectors of **G**
P(•)	:	Probability functions of •
p	:	Process parameters
Q	:	Weighting matrix
r	:	Reference signals
$\mathbf{r}(\boldsymbol{\theta})$	:	Correlation of disturbance and uncertainties
s	:	Complex domain
s_j	:	Slack variables
t	:	Time
U	:	Set of manipulated variable values
u	:	Vector of manipulated variables
W	:	Set of output variable values
w	:	Vector of output variables
$\overline{\mathbf{W}}$	:	Set of dependent variable values
$\overline{\mathbf{w}}$	:	Vector of dependent variables
X	:	Set of state variable values
x	:	Vector of state variables
$\dot{\mathbf{x}}$	:	Vector of state variable derivatives with respect to time
Y	:	Set of discrete design variables
y	:	Vector of discrete design variables
$\mathbf{y^s}$	:	Process in state s
Z	:	Set of continuous design variables
z	:	Vector of continuous design variables
$\overline{\mathbf{z}}$	:	Vector of independent design variables
$\overline{\mathbf{z}}^*$	:	Vector of optimum design variables

Greek letters		
$\times$	:	Element-by-element multiplication
β	:	Closed-loop Disturbance Gain (CLDG)
β_p	:	Partial Disturbance Gain
χ	:	Feasibility function
$\Delta\theta^+$	:	The expected deviations of uncertainties, positive direction
$\Delta\theta^-$	:	The expected deviations of uncertainties, negative direction
Δ_{min}	:	The minimum time delay
δ	:	Deviation due to uncertainties
$\delta(\mathbf{G_1},\mathbf{G_2})$	:	Gap metric between the models $\mathbf{G_1}$ and $\mathbf{G_2}$
$\vec{\delta}(\mathbf{G_1},\mathbf{G_2})$	:	Directed gap from $\mathbf{G_1}$ to $\mathbf{G_2}$
δ_i	:	The amplitude of the i^{th} slack variable
Σ	:	The diagonal matrix of the singular values of **G**
Φ	:	Objective function
Φ_0	:	Objective function at the steady state optimum
$\mathbf{\Phi_q}$	:	Set of objective functions
Γ	:	Performance Relative Disturbance Gain
γ	:	Process condition number
$\gamma_\theta(G)$,	:	Disturbance condition number of G
$\boldsymbol{\theta_e}$	:	External disturbances
$\boldsymbol{\theta_h}$	:	Process uncertainties
Λ	:	RGA matrix
λ_i	:	Lagrange multiplier of the i^{th} active constraint
μ	:	Structural singular value
Θ	:	Set of disturbances and uncertainties
θ	:	Vector of disturbances and uncertainties
θ^N	:	Vector of disturbances and uncertainties at nominal values
σ_{max}	:	Maximum singular value
σ_{min}	:	Minimum singular value
τ	:	Time delay
υ	:	Distance to the active constraint
ω_B	:	Bandwidth
ψ	:	Constraint projection function

Superscripts		
+	:	Positive direction
-	:	Negative direction
N	:	Nominal values
*	:	Optimum values
L	:	Lower bounds
U	:	Upper bounds

Subscripts		
B	:	Bandwidth
f	:	Final value
i	:	Index of equality constraints
j	:	Index of inequality constraints
l	:	Left fractional element∞
r	:	Right fractional element

2.12 References

Acevedo, J. and Pistikopoulos, E. N. (1996). "A Parametric MINLP Algorithm for Process Synthesis Problems under Uncertainty", Industrial and Engineering Chemistry Research, 35, 147 - 158.

Ahmed, S. and Sahinidis, N. V. (1998). "Robust Process Planning under Uncertainty", Industrial and Engineering Chemistry Research, 37, 1883 - 1892.

Anderson, B. D. O. and Moore, J. B. (1990). "Optimal Control: Linear Quadratic Method", New Jersey, Englewood Cliffs.

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Bahri, P. A., Bandoni, A. and Romagnoli, J. (1996a). "Effect of Disturbances in Optimizing Control: Steady State Open Loop Backoff Problem", AIChE Journal, 42 (4), 983 - 994.

Bahri, P. A., Bandoni, A. and Romagnoli, J. (1996b). "Operability Assessment in Chemical Plants", Computers and Chemical Engineering, 20, S787 - S792.

Bahri, P. A., Bandoni, J. A. and Romagnoli, J. A. (1997). "Integrated Flexibility and Controllability Analysis in Design of Chemical Processes", AIChE Journal, 43 (4), 997 - 1015.

Bandoni, A. J., Barton, G. W. and Romagnoli, J. A. (1994). "On Optimizing Control and the Effect of Disturbances: Calculation of the Open Loop Back-Off", Computers and Chemical Engineering, 18S (Supplement), S505 - S509.

Bandoni, J. A. and Romagnoli, J. A. (1989). "Uncertainty in Linear Programming", Numerical Application in Mathematics, Brezinski, C., Scientific Publishing Co, IMACS, 635 - 640.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2000). "Flexibility Analysis and Design of Linear Systems by Parametric Programming", AIChE Journal, 46 (2), 335-354.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2002a). "A Case Study in Simultaneous Design and Control Using Rigorous, Mixed-Integer Dynamic Optimization Models", Industrial and Engineering Chemistry Research, 41, 760-778.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2002b). "Flexibility Analysis and Design Using a Parametric Programming Framework", AIChE Journal, 48 (12), 2851-2686.

Bernardo, F. P. and Saraiva, P. M. (1998). "Robust Optimization Framework for Process Parameter and Tolerance Design", AIChE Journal, 44 (9), 2007 - 2017.

Bristol, E. H. (1966). "On a New Measure of Interactions for Multivariable Process Control", IEEE Transactions of on Automatic Control, AC-11, 133-134.

Bryson, A. E. (1999). "Dynamic Optimization", Addison Wesley Longman, Inc.

Cao, Y., Rossiter, D. and Owens, D. (1997). "Input Selection for Disturbance Rejection under Manipulated Variable Constraints", Chemical Engineering Research and Design, 21, S403 - S408.

Cervantes, A. and Biegler, L. T. (1997). "Large-Scale DAE Optimization Using Simultaneous Nonlinear Programming Formulations", AIChE Journal, 44, 1038 - 1050.

Chenery, S. D. (1997). "Process Controllability Analysis Using Linear and Nonlinear Optimization", Ph.D., University of London, University of London, London.

Chenery, S. D. and Walsh, S. (1998). "Process Controllability Analysis Using Linear Programming", Journal of Process Control, 8, 165 - 174.

Chiu, M. S. and Arkun, Y. (1990). "Decentralised Control Structure Selection Based on Integrity Considerations", Industrial and Engineering Chemistry Research, 29, 369 - 373.

Cuthrell, J. E. and Biegler, L. T. (1987). "On the Optimization of Differential Algebraic Process Systems", AIChE Journal, 33, 1257.

Daoutidis, P. and Kravaris, C. (1991). "Inversion and Zero Dynamics in Nonlinear Multivariable Control", AIChE Journal, 37 (4), 527 - 538.

Dimitriadis, V. D. and Pistikopoulos, E. N. (1995). "Flexibility Analysis of Dynamic Systems", Industrial and Engineering Chemistry Research, 34, 4451 - 4462.

Diwekar, U. M. and Rubin, E. S. (1991). "Stochastic Modelling of Processes", Computers and Chemical Engineering, 15, 105.

Dua, V., Bozinis, N. A. and Pistikopoulos, E. N. (2002). "A Multiparametric Programming Approach for Mixed-Integer Quadratic Engineering Problems", Computers & Chemical Engineering, 26 (4-5), 715-733.

Dua, V. and Pistikopoulos, E. N. (1999). "Algorithm for the Solution of Multiparametric Mixed Integer Nonlinear Optimization Problems", Industrial and Engineering Chemistry Research, 38, 3976.

Duran, M. A. and Grossmann, I. E. (1986). "An Outer Approximation Algorithm for a Class of Mixed Integer Nonlinear Programs", Mathematical Programming, 36, 307.

Elliot, T. R. and Luyben, W. L. (1995). "Capacity-Based Economic Approach for the Quantitative Assessment of Process Controllability During the Conceptual Design Stage", Industrial and Engineering Chemistry Research, 34, 3907 - 3915.

Elliot, T. R. and Luyben, W. L. (1997). "Application of the Capacity-Based Economic Approach to an Industrial Scale Process", Industrial and Engineering Chemistry Research, 36, 1727 - 1737.

Epperly, T. G., Ierapetritou, M. G. and Pistikopoulos, E. N. (1997). "On the Global and Efficient Solution of Stochastic Batch Plant Design Problems", Computers and Chemical Engineering, 21, 1411.

Farschman, C. A., Viswanath, K. P. and Ydstie, B. E. (1998). "Process System and Inventory Control", AIChE Journal, 44 (8), 1841 - 1857.

Figueroa, J. L. (2000). "Economic Performance of Variable Structure Control: A Case Study", Computers and Chemical Engineering, 24, 1821 - 1827.

Figueroa, J. L., Bahri, P. A., Bandoni, J. A. and Romagnoli, J. A. (1996). "Economic Impact of Disturbances and Uncertain Parameters in Chemical Processes - a Dynamic Back-Off Analysis", Computers and Chemical Engineering, 4, 453 - 461.

Floudas, C. A. (1999). "Recent Advances in Global Optimization for Process Synthesis, Design and Control: Enclosure of All Solutions", Computers and Chemical Engineering, S963-973.

Floudas, C. A., Gumus, Z. H. and Ierapetritou, M. G. (2001). "Global Optimisation in Design under Uncertainty: Feasibility Test and Flexibility Index Problems", Industrial and Engineering Chemistry Research, 40, 4267 - 4282.

Friedman, Y. and Reklaitis, G. V. (1975). "Flexible Solutions to Linear Programs under Uncertainty: Inequality Constraints", AIChE Journal, 21, 77 - 83.

Galan, O., Romagnoli, J. A., Palazoglu, A. and Arkun, Y. (2003). "Gap Metric Concept and Implications for Multilinear Model-Based Controller Design", Industrial and Engineering Chemistry Research, 42, 2180 2197.

Gannavarapu, C. (1991). "Economic Assessment in the Synthesis of Optimising Control Schemes", Ph.D. Thesis, Department of Chemical Engineering, University of Sydney, Sydney, Australia.

Garcia, C. E. and Morari, M. (1982). "Internal Model Control, 1. A Unifying Review and Some New Results", Industrial and Engineering Chemistry Process Design, 21, 308 - 323.

Geoffrion, A. m. (1972). "Generalized Benders Decomposition", Journal of Optimization Theory and Application, 10 (4), 237-260.

Georgakis, C., Uzturk, D., Subramanian, S. and Vinson, D. R. (2001). "On the Operability of Continuous Processes", 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-6), Chejudo Island, Korea.

Georgiou, A. and Floudas, C. A. (1989). "Structural Analysis and Synthesis of Feasible Control Systems, Theory and Applications", Chemical Engineering Research and Design, 67, 600 - 618.

Grosdidier, P. and Morari, M. (1986). "Interaction Measures for Systems under Decentralised Control", Automatica, 22, 309 - 319.

Grossman, I. E. and Morari, M. (1983). "Operability, Resiliency and Flexibility - Process Design for Changing World", Second International Conference Foundations of Computer Aided Process Design, Snowmass, CO.

Grossmann, I. E. and Floudas, C. A. (1987). "Active Constraint Strategy for Flexibility Analysis in Chemical Processes", Computers and Chemical Engineering, 11, 675 - 693.

Grossmann, I. E., Halemane, K. P. and Swaney, R. E. (1983). "Optimization Strategies for Flexible Chemical Processes", Computers and Chemical Engineering, 7 (4), 439 - 462.

Grossmann, I. E. and Sargent, R. W. H. (1978). "Optimal Design of Chemical Plants with Uncertain Parameters", AIChE Journal, 24 (6), 1021 - 1028.

Haggblom, K. E. (1997). "Control Structure Selection Via Relative Gain Analysis of Partially Controlled Systems", European Control Conference, ECC 97, Brussels, Belgium.

Halemane, K. P. and Grossmann, I. E. (1983). "Optimal Process Design under Uncertainty", AIChE Journal, 29, 425.

Hangos, K. A., Alonso, A. A., Perkins, J. D. and Ydstie, B. E. (1999). "Thermodynamic Approach to the Structural Stability of Process Plants", AIChE Journal, 45 (4), 802 - 816.

Hernandez, S. and Jimenez, A. (1999). "Controllability Analysis of Thermally Coupled Distillation Systems", Industrial and Engineering Chemistry Research, 38, 3957-3963.

Hirschorn, R. M. (1979). "Invertibility of Multivariable Nonlinear Control Systems", IEEE Transactions of on Automatic Control, 24 (6), 855 - 865.

Holt and Morari, M. (1985). "Design of Resilient Processing Plants V & VI", Chemical Engineering Science, 40 (7), 1229 - 1237.

Hopkins, L., Lant, P. and Newell, R. B. (1998). "Output Structural Controllability: A Tool for Integrated Process Design and Control", Journal of Process Control, 8 (1), 57 - 68.

Hovd, M. and Braatz, R. D. (2000). "On the Computation of Disturbance Rejection Measures", IFAC International Symposium on Advanced Control of Chemical Processes, Pisa, Italy.

Hovd, M. and Skogestad, S. (1994). "Pairing Criteria for Decentralized Control of Unstable Plants", Industrial and Engineering Chemistry Research, 33, 2134 - 2139.

Ierapetritou, M. G. (2001). "New Approach for Quantifying Process Feasibility: Convex and 1-D Quasi-Convex Regions", AIChE Journal, 47 (6), 1407 - 1417.

Ierapetritou, M. G., Acevedo, J. and Pistikopoulos, E. N. (1996). "An Optimisation Approach for Process Engineering Problems under Uncertainty", Computers and Chemical Engineering, 20, 703 - 709.

Ierapetritou, M. G. and Pistikopoulos, E. N. (1996). "Batch Plant Design and Operation under Uncertainty", Industrial and Engineering Chemistry Research, 35, 772 - 787.

Kabatek, U. and Swaney, R. E. (1992). "Worst Case Identification in Structured Process Systems", Computers and Chemical Engineering, 16, 1063 - 1071.

Kalman, R. E. (1960). "On the General Theory of Control Systems", First IFAC Congress, Butterworths, Moscow.

Kim, K. J. and Diwekar, U. M. (2002). "Efficient Combinatorial Optimization under Uncertainty: 1. Algorithmic Development", Industrial and Engineering Chemistry Research, 41, 1276 - 1284.

Kocis, G. R. and Grossmann, I. E. (1987). "Relaxation Strategy for the Structural Optimization of Process Flowsheets", Industrial and Engineering Chemistry Research, 26, 1869 - 1880.

Kookos, I. K. and Perkins, J. D. (2003). "On the Efficient Computation of Disturbance Rejection Measures", Computers and Chemical Engineering, 27, 95 - 99.

Kosut, R. L. (2001). "Iterative Adaptive Control: Windsurfing with Confidence", Model Identification and Adaptive Control: From Windsurfing to Telecommunication, Goodwin, G., Springer-Verlag.

Lenhoff, A. M. and Morari, M. (1982). "Design of Resilient Processing Plants - I: Process Design under Consideration of Dynamic Aspects", Chemical Engineering Science, 37, 245 - 258.

Li, C. W. and Feng, Y. K. (1987). "Functional Reproducibility of General Multivariable Analytic Nonlinear Systems", International Journal of Control, 45 (1), 255 - 268.

Liu, M. L. and Sahinidis, N. (1996). "Optimisation in Process Planning under Uncertainty", Industrial and Engineering Chemistry Research, 35, 4154 - 4165.

Logsdon, J. S. and Biegler, L. T. (1989). "Accurate Solution of Differential - Algebraic Optimization Problems", Industrial and Engineering Chemistry Research, 28, 1628 - 1639.

Luyben, W. L. and Floudas, C. A. (1994a). "Analyzing the Interaction of Design and Control - 1. A Multiobjective Framework and Application to Binary Distillation Synthesis", Computers and Chemical Engineering, 18 (10), 933 - 969.

Luyben, W. L. and Floudas, C. A. (1994b). "Analyzing the Interaction of Design and Control - 1. A Reactor - Separator - Recycle System", Computers and Chemical Engineering, 18 (10), 971 - 994.

Lyman, P. R., Luyben, W. L. and Tyreus, B. D. (1996). "Method for Assessing the Effect of Design Parameters on Controllability", Industrial and Engineering Chemistry Research, 35, 3484 - 3497.

Manousiouthakis, V. and Nikolau, M. (1989). "Analysis of Decentralised Control Structures for Nonlinear Systems", AIChE Journal, 35 (4), 549 - 558.

Manousiouthakis, V., Savage, R. and Arkun, Y. (1986). "Synthesis of Decentralised Control Strructures Using the Concept of Block Relative Gain", AIChE Journal, 32, 991 - 1003.

Marlin, T. E. (1995). "Process Control - Designing Processes and Control Systems for Dynamic Performance", McGraw-Hill. Inc.

McAvoy, T. J. (1983). "Interaction Analysis", North Carolina, USA, Research Triangle Park.

Mijares, G., Holland, C. D., McDaniel, R., Dollar, C. R. and Gallun, S. E. (1985). "Analysis and Evaluation of the Relative Gains for Nonlinear Systems", Computers and Chemical Engineering, 9 (1), 61 - 70.

Mohideen, J., Perkins, J. D. and Pistikopoulos, E. N. (1996). "Optimal Design of Dynamic Systems under Uncertainty", AIChE Journal, 42, 2251-2272.

Morari, M. (1983). "Design of Resilient Processing Plants III, a General Framework for the Assessment of Dynamic Resilience", Chemical Engineering Science, 38, 1881 - 1891.

Morari, M. and Skogestad, S. (1985). "Effect of Model Uncertainty on Dynamic Resilience", I. Chem. E. Symposium Series.

Narraway, L. and Perkins, J. (1994). "Selection of Process Control Structure Based on Economics", Computers and Chemical Engineering, 18 (Supplementary), S511-S515.

Narraway, L. T., Perkins, J. D. and Barton, G. W. (1991). "Interaction between Process Design and Process Control: Economic Analysis of Process Dynamics", Journal of Process Control, 1, 243 - 250.

Ngo, T. and Bahri, P. A. (1996). "Process Systems Engineering and Mineral Industries: Flexibility Analysis in a Flotation Circuit", Optimization '96.

Niederlinski, A. (1971). "A Heuristic Approach to the Design of Linear Multivariable Control Systems", Automatica, 7, 691 - 701.

Novak, Z. and Kravanja, Z. (1999). "Mixed-Integer Nonlinear Programming Problem Process Synthesis under Uncertainty by Reduced Dimensional Stochastic Optimization", Industrial and Engineering Chemistry Research, 38, 2680-2698.

Novak, Z. and Kravanja, Z. (2003). "A Strategy for MINLP Synthesis of Flexible and Operable Processes", Fourth International Conference on Foundation of Computer Aided Process Operations (FOCAPO 2003), Coral Springs, Florida, USA.

Ostrovsky, G. M., Volin, Y. M. and Golovaskhin, D. V. (1998). "Optimization Problem of Complex System under Uncertainty", Computers and Chemical Engineering, 22 (7 - 8), 1007 - 1015.

Palazoglu, A. and Arkun, Y. (1986). "Multiobjective Approach to Design Chemical Plants with Robust Dynamic Operability Characteristics", Computers and Chemical Engineering, 10 (6), 567 - 575.

Palazoglu, A. and Arkun, Y. (1987). "Design of Chemical Plants with Multiregime Capabilities and Robust Dynamic Operability Characteristics", Computers and Chemical Engineering, 11, 205-216.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994a). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Dynamic and Control Structure Consideration", Industrial and Engineering Chemistry Research, 33, 1738 - 1755.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994b). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Flexibility and Structural Controllability Aspects", Industrial and Engineering Chemistry Research, 33, 1718 - 1737.

Perkins, J. D., Gannavarapu, C. and Barton, G. W. (1989). "Simulation, Optimization and Control System Design for an Industrial Process", Computational and Integrated Process Engineering, 114, 141 - 156.

Perkins, J. D. and Wong, M. P. F. (1985). "Assessing Controllability of Chemical Plants", Chemical Engineering Research and Design, 63, 358 - 362.

Pistikopoulos, E. N. and Grossmann, I. E. (1988a). "Evaluation and Redesign for Improving Flexibility in Linear System with Infeasible Nominal Conditions", Computers and Chemical Engineering, 12, 841.

Pistikopoulos, E. N. and Grossmann, I. E. (1988b). "Optimal Retrofit Design for Improving Process Flexibility in Linear System", Computers and Chemical Engineering, 12, 719 - 731.

Pistikopoulos, E. N. and Grossmann, I. E. (1988c). "Optimal Retrofit Design for Improving Process Flexibility in Nonlinear System I. Fixed Degree of Flexibility", Computers and Chemical Engineering, 13, 1003.

Pistikopoulos, E. N. and Grossmann, I. E. (1988d). "Optimal Retrofit Design for Improving Process Flexibility in Nonlinear System II. Optimal Level of Flexibility", Computers and Chemical Engineering, 13, 1087.

Pistikopoulos, E. N. and Grossmann, I. E. (1988e). "Stochastic Optimization of Flexibility in Retrofit Design of Linear Systems", Computers and Chemical Engineering, 12 (12), 1215 - 1227.

Pistikopoulos, E. N. and Mazzucchi, T. A. (1990). "A Novel Flexibility Analysis Approach for Processes with Stochastic Parameters", Computers and Chemical Engineering, 14, 991 - 1000.

Raspanti, C. G., Bandoni, J. A. and Biegler, L. T. (2000). "New Strategies for Flexibility Analysis and Design under Uncertainty", Computers and Chemical Engineering, 24, 2193 - 2209.

Rooney, W. C. and Biegler, L. T. (2003). "Optimal Process Design with Model Parameter Uncertainty and Process Variability", AIChE Journal, 49 (2), 438 - 449.

Rosenbrock (1970). "State-Space and Multivariable Theory", London, UK, Nelson.

Saboo, A. K., Morari, M. and Woodcock, D. C. (1985). "Design of Resilient Processing Plants-VIII: A Resilience Index for Heat Exchanger Networks", Chemical Engineering Science, 40, 1553 - 1565.

Safonov, M. G. (1996). "Focusing on the Knowable", Control Using Logic-Based Switching, Morse, A. S., Berlin, Springer-Verlag, 224 - 233.

Safonov, M. G. and Tsao, T. C. (1997). "The Unfalsified Control Concept and Learning", IEEE Transactions of on Automatic Control, 42 (6), 843 - 847.

Samsatli, N. J., Papageorgiou, L. G. and Shah, N. (1998). "Robustness Metrics for Dynamic Optimisation Models under Parameter Uncertainty", AIChE Journal, 44 (9), 1993 - 2006.

Samyudia, Y., Kadiman, K., Lee, P. L. and Cameron, I. T. (2000). "Gap Metric Based Control of Process with Recycle Streams", International Symposium on Advanced Control of Chemical Processes (ADCHEM 2000), Pisa, Italy.

Samyudia, Y. and Lee, P. L. (2002). "Iterative Design of Robust Generic Model Controller for Industrial Processes", Control Engineering Practice.

Schweiger, C. A. and Floudas, C. A. (1997). "Interaction of Design and Control: Optimization with Dynamic Models", Optimal Control: Theory, Algorithms and Applications, Hager, W. W. and Pardalos, P. M., Kluwer Academic Publishers B. V.

Schweiger, C. A. and Floudas, C. A. (1998). "Process Synthesis, Design and Control: A Mixed-Integer Optimal Control Framework", IFAC Conference on Dynamics and Control of Process Systems, Corfu, Greece, Elsevier.

Singh, S. N. (1982). "Functional Reproducibility of Multivariable Nonlinear Systems", IEEE Transactions of on Automatic Control, 27 (1), 270-272.

Skogestad, S. and Morari, M. (1987). "Effects of Disturbance Directions on Closed-Loop Performance", Industrial and Engineering Chemistry Process Design, 26, 2029-2035.

Skogestad, S. and Wolff, E. A. (1992). "Controllability Measures for Disturbance Rejection", IFAC workshop on Interactions between Process Design and Process Control, London.

Stanley, G., Marino-Galarraga, M. and McAvoy, T. J. (1985). "Shortcut Operability Analysis: 1. The Relative Disturbance Gain", Industrial and Engineering Chemistry Process Design and Development, 24, 1181 - 1188.

Straub, D. A. and Grossman, I. E. (1990). "Integrated Stochastic Metric of Flexibility for Systems with Discrete State and Continuous Parameter Uncertainties", Computers and Chemical Engineering, 14 (9), 967 - 985.

Straub, D. A. and Grossman, I. E. (1992). "Design Optimization of Stochastic Flexibility", Computers and Chemical Engineering, 17, 339 -.

Swaney, R. E. and Grossman, I. E. (1985a). "An Index for Operational Flexibility in Chemical Process Design. Part I: Formulation and Theory", AIChE Journal, 31, 621 - 630.

Swaney, R. E. and Grossman, I. E. (1985b). "An Index for Operational Flexibility in Chemical Process Design. Part II: Computational Algorithm", AIChE Journal, 31, 631 - 641.

Tjoa, I. B. and Biegler, L. T. (1991). "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential - Algebraic Equation Systems", Industrial and Engineering Chemistry Research, 30, 376 - 385.

Trierweiler, J. O. and Engell, S. (1997). "The Robust Performance Number: A New Tool for Control Structure Design", Chemical Engineering Research and Design, 21, S409 - S414.

Tsinias, J. and Kalouptsidis, N. (1983). "Invertibility of Nonlinear Analytic Single Input Systems", IEEE Transactions of on Automatic Control, 28 (9), 931-933.

Varvarezos, D. K., Grossman, I. E. and Biegler, L. T. (1995). "A Sensitivity Based Approach for Flexibility Analysis and Design of Linear Process Systems", Computers and Chemical Engineering, 19, 1301 - 1316.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Veres, S. M., Kuntsevitch, A. V., Hermsmeyer, S. and Wall, D. S. (1995). "Using GBT Version 5.1 in Identification and Control", 1995 MATLAB Conference.

Villadsen, J. and Michelsen, M. L. (1978). "Solution of Differential Equation Models by Polynomial Approximation", Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

Vinson, D. R. and Georgakis, C. (1998). "A New Measure of Process Output Controllability", 5th IFAC Symposium on Dynamics and Control of Process Systems, Corfu, Greece.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10 (2), 185 - 194.

Viswanathan, J. and Grossmann, I. E. (1990). "A Combined Penalty Function and Outer Approximation Method for MINLP Optimisation", Computers and Chemical Engineering, 14, 769-782.

Vu, T. T. L., Bahri, P. A. and Romagnoli, J. (1997). "Operability Considerations in Chemical Processes: A Switchability Analysis", Computers and Chemical Engineering, 21 (4), S143 - S148.

Walsh, S. and Perkins, J. D. (1996). "Operability and Control in Process Synthesis and Design", Advances in Chemical Engineering, Anderson, J. L., Academic Press, 23, 301 - 402.

Weitz, O. and Lewin, D. R. (1996). "Dynamic Controllability and Resiliency Diagnosis Using Steady State Process Flowsheet Data", Computers and Chemical Engineering, 20 (4), 325 - 335.

Wolff, E. A., Perkins, J. D. and Skogestad, S. (1992a). "A Procedure for Operability Analysis", ESCAPE 4.

Wolff, E. A., Skogestad, S., Hovd, M. and Mathisen, K. W. (1992b). "A Procedure for Controllability Analysis", IFAC workshop on Interactions between Process Design and Process Control.

Ydstie, B. E. and Alonso, A. A. (1997). "Process Systems and Passivity Via the Clausius - Planck Inequality", Systems & Control Letters, 30, 253 - 264.

Zhao, Y. and Skogestad, S. (1997). "Comparison of Various Control Configuration for Continuous Bioreactors", Industrial and Engineering Chemistry Research, 36, 697-705.

Zheng, A. and Mahajanam, R. V. (1999). "A Qualitative Controllability Index", Industrial and Engineering Chemistry Research, 38, 999-1006.

Ziegler, J. G. and Nichols, N. B. (1943). "Process Lags in Automatic Control Circuits", Transactions of ASME, 65, 433.

Chapter 3

Integration of the Output Controllability Index within the Dynamic Operability Framework

3.1 Introduction

This chapter presents the formal integration of a controllability index within the Dynamic Operability Framework (Bahri, 1996). The Output Controllability Index (Vinson and Georgakis, 2000) approach is adapted for this purpose, specifically for controllability assessment of regulatory and fixed structure processes. The main added features are the geometric representations of the feasible operating range, input and output spaces in terms of their convex-hull, as well as the corresponding geometric operations. In addition to deliver the best feasible operating conditions, the modified framework also provides a controllability index involving all of the output variables (Generalised Integral Absolute Error, GIAE) and the variation of profit related to disturbance rejection dynamics. The applicability of the modified framework is demonstrated on a nonlinear chemical process system with a fixed control structure.

The early results of this study have been presented as a paper titled "Adaptation of Output Controllability Index within the Dynamic Operability Framework" at the 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS 2001), Jejudo Island, Korea (Ekawati and Bahri, 2001). The completed version of this chapter has been published in the Journal of Process Control, titled "Integration of the Output Controllability Index within The Dynamic Operability Framework in Process System Design" (Ekawati and Bahri, 2003).

The chapter is structured as follows: Section 3.2 discusses the features and the formulation of the original Dynamic Operability Framework (DOF) in facilitating process controllability assessments. Section 3.3 follows with the review, prospects and adaptation of the Output Controllability Index (OCI). Section 3.4 presents the proposed extension of the framework and discusses the numerical strategies as well as its position in relation to parallel works in flexibility and controllability studies. Section 3.5 demonstrates the application of the proposed framework on a nonlinear dynamic chemical process. Finally, section 3.6 summarises and concludes the chapter.

3.2 Dynamic Operability Framework

3.2.1 Features

This section discusses the features of the original Dynamic Operability Framework (Bahri, 1996), particularly those that facilitate the integration of a formal controllability assessment.

The principles of the original approach for a fixed structure are illustrated in Figure 3.1. The feasible operating condition is the area bounded by process constraints. It is typical that the nominal optimum operating point lies on at least one of the constraints. However, this raises concerns about operational feasibility in the presence of disturbances or uncertainties. To address this problem, the nominal operating point is moved inside the feasible region and an operating domain surrounding the point is

defined. This domain represents all possible deviations from the nominal point due to disturbances and uncertainties. Whenever the operating domain lies entirely within the feasible operating condition, it is stated that the process is flexible. It is also typical to ensure that the entire domain is located at the best possible profit, which is as close as possible to the nominal steady state optimum. These efforts result in a design that is optimally flexible and economically viable (Bahri *et al.*, 1996).

It is interesting to note from the above description that the flexibility requirements are the upper bounds of controllability assessments. Process flexibility focuses on either the overshoots or final states of process dynamics. If the whole process dynamics can be captured and assessed, the concept becomes the dynamic flexibility or controllability assessment. In the regulatory case, it can be translated to the disturbance rejection capability. Therefore, it is possible to extend the flexibility study to assess specific problems in controllability, such as disturbance rejection. This prospect is examined further in the framework formulation.

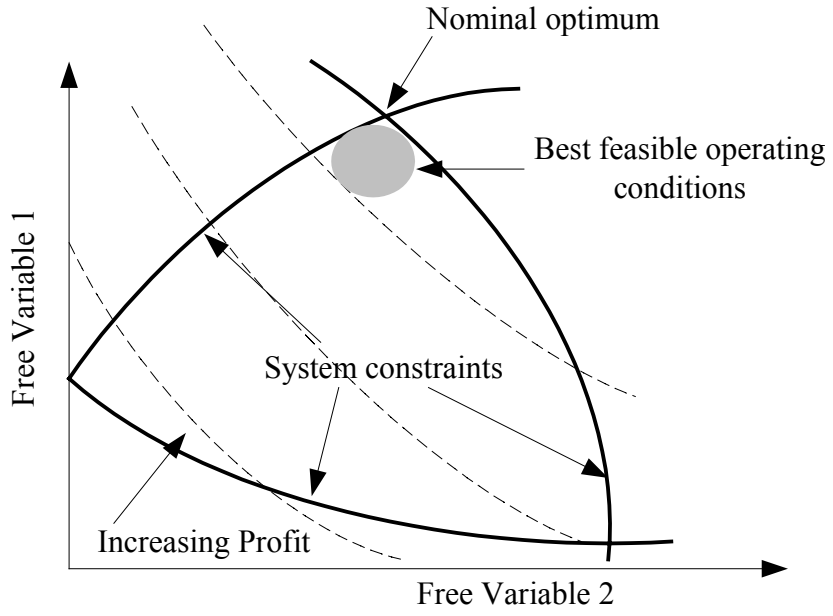


Figure 3.1 The Basic Principles of the Dynamic Operability Framework

3.2.2 Problem Formulation

The implementation of DOF leads to the characterisation of the operating domain and the assessment of its feasibility. The framework formulation is as follows (Bahri, 1996):

$$\min_{\mathbf{z}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{N}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{3.1}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{k}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 \qquad \mathbf{i} \in \mathbf{E}$$

$$\max_{\boldsymbol{\theta}} \max_{\mathbf{j}} \mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{k}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le 0 \qquad \mathbf{j} \in \mathbf{I}$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{x}_{ss}, \mathbf{u}_{ss}, \mathbf{w}_{ss}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z}^{l} \le \mathbf{z} \le \mathbf{z}^{u}\}$$

$$\mathbf{u}(t) \in \mathbf{U} = \{\mathbf{u} : \mathbf{u}^{l} \le \mathbf{u} \le \mathbf{u}^{u}\}$$

$$\boldsymbol{\theta}^{N}, \boldsymbol{\theta}^{k} \in \boldsymbol{\Theta} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{l} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{u}\}$$

$$\mathbf{x} \in \mathbf{X} = \{\mathbf{x} : \mathbf{x}^{l} \le \mathbf{x} \le \mathbf{x}^{u}\}$$

$$\mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w}^{l} \le \mathbf{w} \le \mathbf{w}^{u}\}$$

$$t \in \{t : t_0 \le t \le t_f\}$$

$$\mathbf{x_{ss}} = \mathbf{x}(t = t_0)$$

$$\mathbf{u_{ss}} = \mathbf{u}(t = t_0)$$

$$\mathbf{w_{ss}} = \mathbf{w}(t = t_0)$$

Here, Φ is the objective function. The vectors **x, u** and **w** are the state, manipulated and output variables respectively. The augmented vector $\bar{\mathbf{z}}$ consists of the design variables **z** and the initial conditions $\mathbf{x_{ss}}$, $\mathbf{u_{ss}}$ and $\mathbf{w_{ss}}$. The vector $\boldsymbol{\theta}^{N}$ is the nominal disturbance and uncertainty values, and $\boldsymbol{\theta}^{k}$ is the critical combinations. The augmented vector $\bar{\mathbf{p}}$ consists of process and controller parameters, $\mathbf{p_h}$ and $\mathbf{p_c}$ respectively. The equality constraints $\mathbf{h_i}$ represent the process control model and the inequalities $\mathbf{g_j}$ define the feasible operating region. The set **E** and **I** are the indices of equality and inequality constraints, respectively. The sets **Z**, **U**, Θ, **X** and **W** contain the possible values of the respective variables. The set Θ is the set of step functions with uniformly distributed magnitudes. The variable t is time, where t_o is the initial time, t_f is the final time, and $\bar{t}$ is the specific time to assess the objective function.

The formulation utilises the controller parameterisation approach. That is, the controller parameters $\mathbf{p_u}$ are designed using any control methods. Then, further tuning is realised by assigning the scaling coefficients for $\mathbf{p_u}$ as one of the decision variables.

It is apparent that equation (3.1) is a Non-Linear Semi-Infinite Dynamic Programming (NLSIDP) problem. Bahri (1996) formulated an iterative dynamic optimisation algorithm consisting of two levels of optimisation, which are the outer-level and the inner-level; to solve this problem. The outer-level returns the optimal design variables through the solutions of dynamic Non-Linear Programming (NLP), and the inner-level investigates the feasibility of the operating conditions found in the preceding outer-level. The iteration continues, until the operating conditions are completely feasible. The mathematical formulation for this two-level optimisation is as follows:

Outer-level:

$$\min_{\mathbf{z}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{3.2}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 \qquad \mathbf{i} \in \mathbf{E}$$

$$\max_{\boldsymbol{\theta}} \max_{\mathbf{j}} \mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le 0 \qquad \mathbf{j} \in \mathbf{I}$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{x_{ss}}, \mathbf{u_{ss}}, \mathbf{w_{ss}}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z^l} \le \mathbf{z} \le \mathbf{z^u}\}$$

$$\boldsymbol{\theta}^{\mathbf{N}}, \boldsymbol{\theta}^{\mathbf{k}} \in \boldsymbol{\Theta} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathrm{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathrm{u}}\},$$

$$\mathbf{x}(t) \in \mathbf{X} = \{\mathbf{x} : \mathbf{x^l} \le \mathbf{x} \le \mathbf{x^u}\}$$

$$\mathbf{u}(t) \in \mathbf{U} = \{\mathbf{u} : \mathbf{u^l} \le \mathbf{u} \le \mathbf{u^u}\}$$

$$\mathbf{w}(t) \in \mathbf{W} = \{\mathbf{w} : \mathbf{w^l} \le \mathbf{w} \le \mathbf{w^u}\}$$

$$t \in \{t : t_o \le t \le t_f\}$$

$$\mathbf{x_{ss}} = \mathbf{x}(t = t_o)$$

$$\mathbf{u_{ss}} = \mathbf{u}(t = t_o)$$

$$\mathbf{w_{ss}} = \mathbf{w}(t = t_o)$$

Inner-level:

$$\max_{\boldsymbol{\theta}} \mathbf{g_j}(\bar{\mathbf{z}}^*, \boldsymbol{\theta}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \bar{\mathbf{p}}, t) \tag{3.3}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}^*, \boldsymbol{\theta}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \bar{\mathbf{p}}, t) = 0 \qquad \mathbf{i} \in \mathbf{E}$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{x_{ss}}, \mathbf{u_{ss}}, \mathbf{w_{ss}}\}$$

$$\boldsymbol{\theta} \in \boldsymbol{\Theta} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^l \le \boldsymbol{\theta} \le \boldsymbol{\theta}^u\}$$

$$t \in \{t : t_o \le t \le t_f\}$$

$$\mathbf{x_{ss}} = \mathbf{x}(t = t_o)$$

$$\mathbf{u_{ss}} = \mathbf{u}(t = t_o)$$

$$\mathbf{w_{ss}} = \mathbf{w}(t = t_o)$$

Vector $\bar{\mathbf{z}}^*$ contains the optimal values found in the preceding outer-level. The vector, $\boldsymbol{\theta}^k$ is the combination of disturbances/uncertainties found in the preceding inner-levels, and the rest of variables are similar to (3.1).

The initial outer-level optimises the process with respect to the nominal disturbances and uncertainties $\boldsymbol{\theta}^N$, hence a steady state optimisation. The optimum solution $\bar{\mathbf{z}}_1^*$ is subsequently supplied to the inner-level. In this level, the expected disturbances and uncertainties are applied to the process. The multiple maximisation problems in (3.3) principally assess the peak of process dynamic profiles against each constraint. If there is no constraint violation in this step, the process is completely feasible and the algorithm terminates. Otherwise, the critical disturbance and uncertainty combinations $\boldsymbol{\theta}^k$, the set that causes the worst violations, is identified and supplied to the outer-level of the next iteration. In this subsequent outer level, both the process nominal steady state involving $\boldsymbol{\theta}^N$, and the process dynamics due to $\boldsymbol{\theta}^k$ are optimised. The result $\bar{\mathbf{z}}_2^*$ is sent to the inner level. The iteration continues until no constraint violations found in the corresponding inner level.

The peak profiles found in the inner-level can be interpreted as the maximum absolute errors that determine the upper bound of process controllability. As discussed in Chapter 2, this is a very limited quantification of controllability. A better measurement can be obtained if the observation is extended to the whole profiles. In addition, it

must take accounts of the interaction between process variables. For instance, reducing variation in one variable may cause more variation in other equally important variables, and may affect the overall controllability. Capturing these aspects requires more flexible characterisation of process dynamics. The adaptation of Output Controllability Index (OCI) approach to achieve such task is described in the following section.

3.3 Adaptation of Output Controllability Index

The Output Controllability Index (OCI) approach by Vinson and Georgakis (1998; 2000) are discussed below, followed by its adaptation into the Dynamic Operability Framework of the regulatory cases.

3.3.1 Definitions

Vinson and Georgakis (1998; 2000) introduced their controllability approach using several definitions of variable sets as follows:

1. Available Input Set (AIS): the set of values of the manipulated variables.
2. Desired Output Space (DOS): the set of desired values of the output variables.
3. Achievable Output Space (AOS): the set of the output values that could be achieved based on the available values of manipulated variables.
4. Desired Input Space (DIS): the set of values of the manipulated variables required to achieve the DOS.
5. Expected Disturbance Space (EDS): the set of the disturbances and uncertainties expected to affect the process.

The approach characterises these sets as multi-dimensional polyhedral spaces. The AIS and DOS are defined by the constraints of the feasible input and output spaces respectively. Subsequently, AOS is the solution of the steady state process model for the entire AIS. Based on the present nomenclature, one can say that the set member $\mathbf{w}$ of AOS is defined by $\mathbf{w}=\mathbf{h}(\mathbf{u})$.On the other hand, DIS is the solution of the inverse of process model for the entire DOS, where the set member $\mathbf{u}$ is defined by $\mathbf{u}=\mathbf{h}^{-1}(\mathbf{w})$.

Having outlined these spaces, several variations of Output Controllability Index (OCI) are introduced. Firstly, the servo-Output Controllability Index (OCI) was defined in the output space, as follows (Vinson and Georgakis, 2000):

$$\text{s - OCI} = \frac{\mu(\text{AOS}_u \cap \text{DOS})}{\mu(\text{DOS})} \tag{3.4}$$

Here, μ is a function to calculate the size of the respective space. For instance, μ is an area in two-dimensional space, or a volume in three-dimensional space. The index s-OCI quantifies the ability to achieve desired outputs at the nominal disturbances and uncertainties condition. A small index implies lower process controllability. This concept is illustrated for two-dimensional linear process in Figure 3.2, where the shaded area is the intersection between AOS and DOS.

The servo-OCI was also defined in the input space using the inverse of process transfer function as follows (Vinson and Georgakis, 2000):

$$\text{s - OCI} = \frac{\mu(\text{AIS} \cap \text{DIS}_w)}{\mu(\text{DIS}_w)} \tag{3.5}$$

Here, the member of DIS_w is calculated by $\mathbf{u=h^{-1}(w)}$.

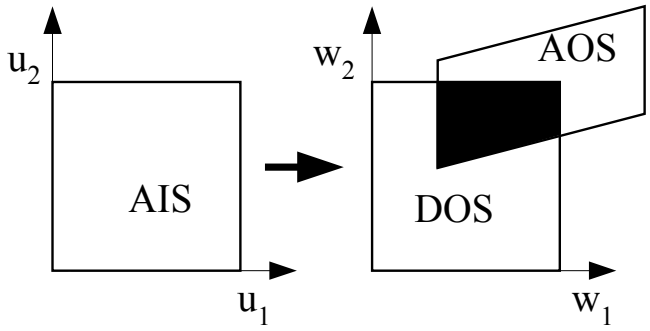


Figure 3.2 The Original Concept of r-OCI in the Output Space

The ability to compensate for the effects of disturbances was addressed in the regulatory-OCI (Vinson and Georgakis, 2000). For this purpose, EDS and DOS were transformed back to the required input space through the inverse **w=h(u,θ)** in terms of **u**. At the nominal and fixed values of **w** and with $\theta \in$ EDS, the solution is DIS_θ. Hence, the steady state regulatory–OCI, which quantifies the capability to compensate the EDS and to maintain the outputs at nominal values, is defined as follows:

$$r\text{-}OCI = \frac{\mu(AIS \cap DIS_\theta)}{\mu(DIS_\theta)} \tag{3.6}$$

When the above task is expanded with different desired values of output variables, the overall DIS is considered. This space is the union of the spaces of the servo (DIS_w) and regulatory (DIS_θ) tasks as follows:

$$DIS = \bigcup_{\theta \in EDS} DIS_w(\theta) = \bigcup_{y \in DOS} DIS_\theta(w) \tag{3.7}$$

Therefore, the overall OCI of the process is as follows (Vinson and Georgakis, 2000):

$$OCI = \frac{\mu(AIS \cap DIS)}{\mu(DIS)} \tag{3.8}$$

The calculation of OCI for general and multi-dimensional process involves the construction, the intersection and the ratio between the polyhedral spaces. The construction of projected spaces such as AOS and DIS starts with discretising the perimeter of the respective input spaces into grids. Subsequently, the transfer function at each gird point is solved to define the projected shapes. The computation of the intersections and volumes follows. Vinson and Georgakis performed these geometric operations using the Geometric Bounding Toolbox (GBT) (Veres *et al.*, 1995). Veres and coworker themselves actively utilised these characterisations for iterative identification and control of linear systems (Veres, 1999; Veres *et al.*, 1995; Veres and Mayne, 2001; Veres *et al.*, 1999). Further discussions about the operations and alternative numerical algorithms and software are given in section 3.4.2.2.

3.3.2 Prospects of Integration

The OCI (Vinson and Georgakis, 1998; 2000) quantifies the controllability of multi-inputs and multi-outputs (MIMO) process in a single number. Because it is a ratio of similar quantities, OCI is dimensionless, with value between one, for the best controllability, to zero, for uncontrollable process. For two-dimensional linear and nonlinear processes, it was shown that the index was sensitive to the changes in both process and disturbance gains, in contrast to the Relative Gain Analysis and the Closed-loop Disturbance Gain. The index was also free from scaling problems inherent to the Singular Value analysis and the Condition Number.

Those features pose OCI as an attractive quantification of process flexibility or controllability. However, the computation of index in input space (3.8) limits the application to process models with unique inverse; which is not generally available in chemical processes. Furthermore, facing the constrained resources and goals, which are represented by manipulated and output variables respectively, the index in input space provides pessimistic views by focusing on the limitation of achieving the goals and suggests increasing the resources. The output space index in equation (3.4), on the other hand, does not require model inversion, hence applicable for general nonlinear process models. It also provides optimistic view in terms of expectation of the achievable goals based on the available resources.

The analogy of the variable set with multi-dimensional geometry spaces, is well known for its simple and intuitive appeal (Berg *et al.*, 1997). In flexibility studies, Halemane and Grossman (1981a,c in Grossmann *et al.*, 1983) show that the solutions of linear and jointly convex problems are the functions of critical variables located at the vertices of the polyhedron defining the parameter space. However, numerical tools for multi-dimensional geometry operations with acceptable computational cost, especially for use within optimisation frameworks, were unavailable at the time. Therefore, instead of geometric computations, the problem was solved using Linear Programming (LP) procedure.

The geometric computations indeed have been existed since the ancient Greeks. However, the studies of multi-dimensional geometry computation, with significant interest on efficient algorithm with reduced the computational costs, were only initiated in the 1970s in the field of computational geometry. It is accommodated by the growth of computer graphics, with Ph.D. thesis of Shamos (1978) being one of the milestones (Berg *et al.*, 1997). Over the next twenty-five years, a significant body of useful geometric computation algorithms and software have been produced, including those required to determine OCI (Barber *et al.*, 1996; Veres and Mayne, 2001).

Based on these developments, it is considered the right time to utilise the available computational geometry tools to provide a controllability index within an optimisation framework. The prospect is explored in this study. In the next section, attention is focused on the adaptation of servo-OCI in output space (3.4), for integration into the Dynamic Operability Framework.

3.3.3 Adaptation of the Regulatory OCI

In regulatory cases, the Dynamic Operability Framework focuses on the effect of disturbances and uncertainties on process outputs. This can be translated into the relationships between EDS, AOS and DOS. Suppose the AOS captures the disturbance effects on process outputs in the regulatory case. Hence, it is computed based on the values of EDS and its notation is modified to AOS_θ. Consider as well DOS as the representation of the feasible operating range of the process. If AOS_θ is completely inside the DOS, then the process is completely flexible. This flexibility condition is represented efficiently by the ratio between the size of an intersection of AOS_θ with DOS, and the size of AOS_θ. Therefore, the definition of OCI for the regulatory case (r-OCI) can be adapted as follows:

$$\text{r}-\text{OCI} = \frac{\mu(\text{AOS}_\theta \cap \text{DOS})}{\mu(\text{AOS}_\theta)} \tag{3.9}$$

Here, r-OCI = 1 implies that the process is completely flexible and controllable, since it enforces all of AOS_θ inside the DOS.

The original steady state concept of OCI is also expanded to capture process dynamics. In this adaptation, time becomes an optional dimension of AOS_θ space. The exclusion leads to the flexibility case, where AOS_θ captures the peak responses in output space. Since it involves all of the dynamic output variables, the size of AOS_θ can be considered as the approximation of the Generalised Absolute Error (GAE) of the process. This is the generalisation of the Integral Absolute Error (IAE) concept to the multi-dimensional space. The GAE is formulated as follows:

$$\text{GAE} = \left\| \mathbf{w} - \mathbf{w}_{ss} \right\|_1 \tag{3.10}$$

Accordingly, the inclusion of time in AOS_θ provides the approximation of the Generalised Integral Absolute Error (GIAE) as follows:

$$\text{GIAE} = \int_{t=t_0}^{t_f} \left\| \mathbf{w} - \mathbf{w}_{ss} \right\|_1 dt \tag{3.11}$$

These indices provide broader view to multivariable process dynamics. Therefore these are proposed as alternatives to the conventional and single-function of time indices, such as the Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE) or Integral Squared Error (ISE). The adapted concept is illustrated in Figure 3.3, where the shaded space is the intersection between DOS and AOS_θ.

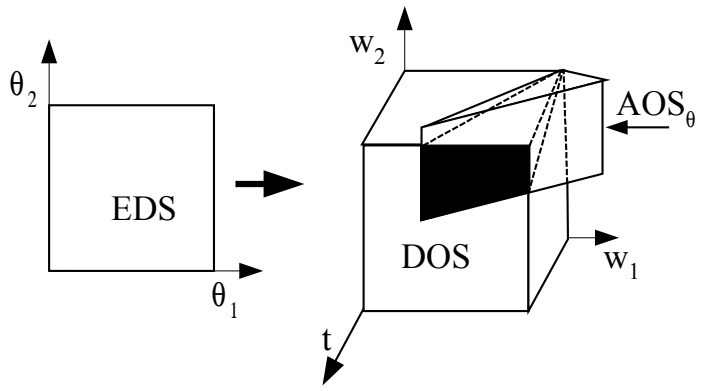


Figure 3.3 The Adapted Concept of r-OCI in the Output Space

3.4 The Proposed Framework

3.4.1 Formulation

The main objective of the feasibility test in the inner-level of the original Dynamic Operability Framework is to verify that the current optimal solutions $\bar{\mathbf{z}}^*$ does not violates the inequality constraints $\mathbf{g_j}(\bar{\mathbf{z}}^*, \boldsymbol{\theta}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)$. To accommodate the adaptation of r-OCI within the framework, these constraints are restructured. The constraints involving the dynamic variables this time are represented by n_g-dimensional polyhedron DOS. Here, n_g is the number of output/measurement variables plus one, where the additional dimension is time ($\mathbf{t}$). The EDS is represented by an n_θ-dimensional polyhedron, where n_θ is the number of disturbance variables. Then, AOS_θ is the projection of EDS into n_g-dimensional output space. The size of AOS_θ, in terms of the volume of the n_g-dimensional space, represents the process controllability.

The integration of the adapted r-OCI into the framework is formulated as follows:

$$\min_{\bar{\mathbf{z}}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{3.12}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = \mathbf{0} \qquad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t)) \le \mathbf{0} \qquad \mathbf{j} \in \mathbf{I_r}$$

$$\mathbf{r-OCI} = \frac{\max_{\boldsymbol{\theta}} \mathbf{AOS_\theta}(\bar{\mathbf{w}}(t))_{t_0 \le t \le t_f} \cap \mathbf{DOS}(\bar{\mathbf{w}})}{\mathbf{DOS}(\bar{\mathbf{w}})} = \mathbf{1}$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{x}_{ss}, \mathbf{u}_{ss}, \mathbf{w}_{ss}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z}^{\mathbf{l}} \le \mathbf{z} \le \mathbf{z}^{\mathbf{u}}\}$$

$$\boldsymbol{\theta}^{\mathbf{N}}, \boldsymbol{\theta}^{\mathbf{k}} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\}$$

$$\bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}$$

$$\mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w}^{\mathbf{l}} \le \mathbf{w} \le \mathbf{w}^{\mathbf{u}}\}$$

$$\mathbf{x}_{ss} = \mathbf{x}(t = t_o)$$

$$\mathbf{u}_{ss} = \mathbf{u}(t = t_o)$$

$$\mathbf{w}_{ss} = \mathbf{w}(t = t_o)$$

$$t \in \{t : t_o \le t \le t_f\}$$

Similar to the original framework, the above problem is formulated in a two-level algorithm as follows:

Outer-level:

$$\min_{\bar{\mathbf{z}}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{3.13}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 \qquad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{k}(t)) \le 0 \qquad \mathbf{j} \in \mathbf{I}_r$$

$$\mathbf{r-OCI} = \frac{\mathbf{AOS}_{\boldsymbol{\theta}}(\bar{\mathbf{w}}(t))_{t_o \le t \le t_f} \cap \mathbf{DOS}(\bar{\mathbf{w}})}{\mathbf{DOS}(\bar{\mathbf{w}})} = 1$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{x}_{ss}, \mathbf{u}_{ss}, \mathbf{w}_{ss}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z}^{\mathbf{l}} \le \mathbf{z} \le \mathbf{z}^{\mathbf{u}}\}$$

$$\boldsymbol{\theta}^{\mathbf{N}}, \boldsymbol{\theta}^{\mathbf{k}} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\}$$

$$\bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}$$

$$\mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w}^{\mathbf{l}} \le \mathbf{w} \le \mathbf{w}^{\mathbf{u}}\}$$

$$\mathbf{x}_{ss} = \mathbf{x}(t = t_o)$$

$$\mathbf{u}_{ss} = \mathbf{u}(t = t_o)$$

$$\mathbf{w}_{ss} = \mathbf{w}(t = t_o)$$

$$t = \{t : t_o \le t \le t_f\}$$

Inner-level:

$$\max_{\boldsymbol{\theta}} \mathbf{AOS}_{\boldsymbol{\theta}}(\bar{\mathbf{w}}(t)) \tag{3.14}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}^*, \boldsymbol{\theta}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 \quad \mathbf{i} \in \mathbf{E}$$

$$\bar{\mathbf{z}}^* \in \{\mathbf{z}^*, \mathbf{x}_{ss}^*, \mathbf{u}_{ss}^*, \mathbf{w}_{ss}^*\}$$

$$\bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}$$

$$\boldsymbol{\theta} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\}$$

$$\mathbf{x_{ss}}^* = \mathbf{x}(t = t_o)$$

$$\mathbf{u_{ss}}^* = \mathbf{u}(t = t_o)$$

$$\mathbf{w_{ss}}^* = \mathbf{w}(t = t_o)$$

$$t \in (t : t_o \le t \le t_f\}$$

In the proposed framework, the inequalities $\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}, \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le \mathbf{0}$, $\mathbf{j} \in \mathbf{I}$ in (3.1) have been reduced into $\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t)) \le \mathbf{0}, \mathbf{j} \in \mathbf{I}_r$. The reduced set involves only the fixed design variables $\bar{\mathbf{z}}$. Meanwhile, the dynamic profiles this time are augmented in the matrix $\bar{\mathbf{w}}(t)$, which subsequently defines the DOS, AOS_θ, and r-OCI. The new constraint r-OCI = 1 enforces the complete feasibility of the dynamic profiles, while the objective function Φ pushes the AOS_θ towards of the optimum profit.

Similar to the original framework, the initial outer-level in the proposed approach is a steady state process optimisation under nominal operating conditions. There is no calculation of AOS_θ, DOS, and r-OCI at this stage. The optimal operating conditions $\bar{\mathbf{z}}_1^*$ are then supplied to the inner-level for feasibility assessment.

Since $\bar{\mathbf{z}}_1^*$ is fixed in the inner-level, only AOS_θ is considered. Instead of performing multiple maximisation as in (3.3), the critical combination $\boldsymbol{\theta}^{\mathbf{k}}$ is determined by exploiting the geometrical feature of the AOS_θ. Here, EDS is discretised into uniform grids, and the process dynamics due to these points are simulated and recorded. The collection of these profiles constructs the polyhedron AOS_θ. The reason for using all points inside EDS is to anticipate any non-monotonic trends or discontinuities in AOS_θ. In all cases, the outer surface of AOS_θ is constructed by $\boldsymbol{\theta}^{\mathbf{k}}$. This critical combination can be extracted effectively by approximating the surface with the convex-hull of AOS_θ, then projecting the convex-hull vertices to the EDS.

The second outer-level solves the optimisation problem (3.13). This includes both process steady states due to $\boldsymbol{\theta}^{\mathbf{N}}$ and process dynamics due to $\boldsymbol{\theta}^{\mathbf{k}}$. The AOS_θ, DOS, and r-OCI are also calculated from the dynamic profiles due to $\boldsymbol{\theta}^{\mathbf{k}}$. In this level, r-OCI is calculated based on the convex-hull volumes of the spaces involved. The subsequent inner-level justifies the process feasibility when no new $\boldsymbol{\theta}^{\mathbf{k}}$ is found, which is typically the case since the complete set of $\boldsymbol{\theta}^{\mathbf{k}}$ has already been considered in the outer-level.

It is easy to see that r-OCI is the flexibility index, because AOS_θ is formed due to $\boldsymbol{\theta}^k$. The quality of the process controllability is also represented by the size of AOS_θ. If time is excluded from AOS_θ dimensions, the size approximates the General Absolute Error (GAE) of the process. Otherwise, it approximates the General Integral Absolute Error (GIAE). The dynamic cost or profit variation can also be included in $\overline{\mathbf{w}}(t)$.

Finally, the proposed modification of the framework can be summarised as follows:

1. The controllability assessment within operability assessment framework is facilitated through the geometric characterisation of the feasible operating region and the projection of disturbance to output space. Here, the flexibility becomes a special case of the controllability problem.
2. The multiple maximisation problems in the inner-level are replaced by geometric operations and an equality constraint, respectively.

3.4.2 Computational Strategies

The computational implementation of the proposed framework involves two major tasks. These are the dynamic optimisation in the outer-levels, and the geometric calculation of the controllability index. The following discussions show how these tasks interact, affecting the optimisation problem, and influencing the computational strategies.

3.4.2.1 Dynamic Optimisation and Dynamic Solver

Computation of controllability index in the outer-level (3.12) requires the solution of $\mathbf{h_i}(\overline{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0$. This biases the optimisation problem toward the sequential method, and calls for proper dynamic solvers.

The performance of dynamics solvers depends on the index of Differential Algebraic Equation (DAE) representing the process model. Formally, an index of a DAE is defined as the minimum number of differentiation required transforming a DAE to explicit Ordinary Differential Equations (ODE). In practice, the degree-of-freedom of

a DAE are higher than the number of differential equations. Such model requires more initial values than ODE. These extras are implicitly constrained within the model equations, and are not to determine independently (Fabian *et al.*, 2001).

Consider the DAE $\mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0$. In dynamic simulations, the DAE problems arise when the dynamics of manipulated and output variables, $\mathbf{u}(t)$ and $\mathbf{w}(t)$ respectively, are not linear and separable functions of state variable $\mathbf{x}(t)$. In this case, the initial condition $\mathbf{u_{ss}}$ and w_{ss} become the extra degree-of-freedom of the process, in addition to $\mathbf{x_{ss}}$. These initial values strongly determine the accuracy of DAE solution.

The presence of these extra degree-of-freedom in steady state cases has been well known in chemical processes, and many cases have been solved efficiently using commercial Nonlinear Programming (NLP) optimisers. This feature has been fully exploited by simultaneous method to alleviate the DAE index problem. The method discretises the dynamic profiles at a number of points (e.g. using finite element (Logsdon and Biegler, 1989) or orthogonal collocation methods (Tjoa and Biegler, 1991)). Then the values at every point are included in the NLP solution. For a model with n_z design variables and n_w dynamic variables, which is discretised at n_c points, the number of NLP solutions became $n_z+n_w \times n_c$. Hence, the optimisation is converted into large-scale steady state NLP problem (Vassiliadis *et al.*, 1994). The simultaneous method accommodates the path constraints on $\mathbf{u}(t)$ and $\mathbf{w}(t)$. However, the typical size of the problem is huge, even for simple processes. To compensate this problem, the performance of the NLP have been improved using the Sequential Quadratic Programming (SQP) (Tjoa and Biegler, 1991) as well as the reduced-SQP (Cervantes and Biegler, 1997) methods. Nevertheless, the industrial applications are still limited.

Meanwhile, without DAE solver, the sequential methods unable to enforce the path constraints on $\mathbf{u}(t)$ and $\mathbf{w}(t)$. Therefore, the recent ODE and index-1 DAE solver, such as LSODI (Hindmarsh, 1980), SPRINT (Berzins *et al.*, 1989), DASSL (Brenan *et al.*,

1996), and MATLAB ODE Suite (Shampine and Reichelt, 1997), provide the much needed supports for sequential dynamic optimisation method.

Within these solvers, DAE are considered in mass-matrix equations as follows:

$$\begin{aligned} &\mathbf{M}(t,\overline{\mathbf{w}})\times\dot{\overline{\mathbf{w}}}(t)=\mathbf{f_i}(\overline{\mathbf{z}},\boldsymbol{\theta}^{\mathbf{k}}(t),\overline{\mathbf{w}}(t),t) \\ &\overline{\mathbf{z}}\in\{\mathbf{z},\mathbf{x}_0,\mathbf{u}_0,\mathbf{w}_0\} \\ &\overline{\mathbf{w}}(t)\in\{\mathbf{x}(t),\mathbf{u}(t),\mathbf{w}(t)\} \end{aligned} \tag{3.15}$$

A constant, diagonal, and nonsingular mass matrix $\mathbf{M}(t,\overline{\mathbf{w}})$ represents the familiar ODE, while the nonsingular counterpart represents an index-1 DAE. The non-singularity in this case is generated by the algebraic variables **u** and **w**. The dependency on time or other state variables represents higher indices and nonlinearity, which are typically harder to solve. In all cases, the solution requires consistent initial values in $\overline{\mathbf{z}}$ and initial slope $\dot{\overline{\mathbf{w}}}(0)$ such that:

$$\mathbf{M}\times\dot{\overline{\mathbf{w}}}(0)=\mathbf{f_i}(\overline{\mathbf{z}},\boldsymbol{\theta}^{\mathbf{N}}) \tag{3.16}$$

Many aspects of these solvers and related numerical algorithms, such as determination of consistent initial conditions, stiffness, and robustness of integration methods, are still under intensive research and development (Shampine *et al.*, 1997).

Nevertheless, the solvers provide the required dynamic profiles, which include the state variable **x**, the manipulated variable **u** and the measurement variable **w**. The objectives such as cost or performance functions may be integrated in **x** or **w,** each provides the accumulated or instantaneous values, $\mathbf{x}(\bar{\mathbf{t}})$ or $\mathbf{w}(\bar{\mathbf{t}})$ respectively. In process simulation, the proportional and integral control actions are also easily embedded into **u** and **x**. This strategy has wide applications, since many well-known multivariable and nonlinear control methods are realisable in some form of proportional and integral actions. For example, a chemical engineer may consider the Linear Quadratic Gaussian (LQG) controllers as interactive multiple proportional controllers with implicit decoupling actions. Nonlinear controllers such as those designed using Generic Model Control

(GMC) or Input Output linearisation (I/O) method are also realisable in proportional–integral control formats.

The derivative control actions, however, can be challenging. Note that a DAE is a translation of multiple-degrees differential equations into a set of first-degree differential plus algebraic equations. Suppose the derivative control actions are applied on the state variables originated from highest derivatives. These require further model differentiation, leading to the increase of the DAE index, with new state variables and initial conditions. Considering this situation, the process control simulation involving derivative control action must be handled with care. Some form of index reduction, such as substitution strategy (Fabian *et al.*, 2001) may be required.

The proposed Dynamic Operability Framework accommodates the DAE problem. The consistent initial condition is inherently provided within the framework structure and with careful modeling as follows: Firstly, all dynamic variables and degrees-of-freedom **x**, **u** and **w**, should be identified. Then, the steady state and dynamic models are arranged similarly according to the format in (3.15), except that the left-hand side of steady state model is an all-zero vector. In the outer-level (3.13), the steady state model constrains the initial values of the subsequent DAE, as follows:

$$\mathbf{h}_{\mathbf{i}_{\mathbf{ss}}}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}, \mathbf{p}) = \dot{\bar{\mathbf{w}}}(0) \rightarrow 0 \qquad \mathbf{i}_{\mathbf{ss}} = \mathbf{i}_{\mathbf{d}} \in \mathbf{E} \tag{3.17}$$

$$\mathbf{h}_{\mathbf{i}_{\mathbf{d}}}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \bar{\mathbf{w}}(t), \dot{\bar{\mathbf{w}}}(t), \mathbf{p}, t) = 0 \tag{3.18}$$

The steady state model (3.17) naturally forces $\dot{\bar{\mathbf{w}}}(0) = 0$. If utilised in combination with DAE (3.18), as in the second and later framework iterations, the intermediate values $\dot{\bar{\mathbf{w}}}(0)$ on every constraint evaluations become the initial slopes for DAE solution. This arrangement guarantees the solution for index-1 DAE problem, and may solve the higher index counterparts.

Due to the availability of integrated numerical simulations and graphic presentations, the framework is developed in MATLAB. Therefore, the dynamic solver used is the ODE Suite (Shampine and Reichelt, 1997), specifically the function *ode15s.m*. The

requirement for deeper analysis on higher-index DAE is nevertheless acknowledged, and recommended for future development of the framework.

The DAE solver is used along with MATLAB NLP optimiser *fmincon.m*. This optimiser uses either large-scale or medium-scale algorithms. The large-scale algorithm is used if the gradient of objective function available, and if either only upper and lower bounds or linear equality constraints exist. The algorithm uses the subspace trust region based on the interior-reflective Newton method (Coleman and Li, 1996). This is an iterative algorithm, which approximates the solution of a large linear system in each iterations using the Preconditioned Conjugate Gradients (PCG) method. The medium-scale optimization is performed using the Sequential Quadratic Programming (SQP) method. The respective Quadratic Programming (QP) sub-problems are solved at each optimiser-iteration using an active set strategy. The approximation of the Hessian matrix in this sub-problem is updated using the BFGS Quasi-Newton method, which uses a mixed quadratic and cubic line search procedure formula (Mathworks, 2000a).

Due to nonlinearity and nonconvexity inherent to dynamic optimisation problem, this optimiser does not guarantee a global optimum. Therefore, trials with several different initial guesses (multi-start approach) are required to verify the solution. The analysis of finding a global optimum within the proposed framework is considered for future works.

3.4.2.2 OCI Calculation

The above DAE solver provides the required dynamic profiles due to EDS for AOS_θ construction within the proposed framework. The AOS_θ then can be intersected with DOS leading to r-OCI computation in the outer-levels, or mapped back to EDS for critical disturbance θ^k extraction in the inner-levels. These geometric operations involve a representation of the respective multi-dimensional spaces in terms of polytopes, half-spaces and convex-hulls. The available software for two- up to seven-dimensional spaces include the GBT (Veres and Mayne, 2001), Qhull (Barber, 1996),

and CDD (Fukuda, 1996). The following discussion focuses on the required geometric computation, selection of the software and the implementation within the framework.

The pure geometric operations required within the proposed framework include the computation of AOS_θ convex-hull, as well as the intersections of DOS and AOS_θ and its convex-hull. The related and follow-up procedures include the DOS preparation for intersections, discretisation of EDS, and extraction of θ^k using the AOS_θ convex-hull vertices. These operations involve a large number of data. Therefore, it is important to ensure that the geometric solver can handle these efficiently.

The steady state OCI (Vinson and Georgakis, 2000) is computed using GBT, which provides the functions for all of above operations within MATLAB environment. One important difference between their approaches with the proposed framework is that the latter involves much higher numbers of data points in OCI computation. For example, consider a two–dimensional process where each input is discretised into six grids. Vinson and Georgakis step along the boundaries to project the input into output spaces using a steady state model, yielding 21 non-duplicate points in the output space. In contrast, the proposed framework uses all the points in the input space (36 points), and collects the corresponding dynamic profiles. For instance, for 100 seconds with one-second intervals, there would be 3636 data points for each output variable.

GBT uses recursive polytope updating, where each additional point is assessed against all other checked points. As the number of data grows, the computing time increases almost exponentially. Therefore, this software is not suitable for dynamic operability computations, and the alternatives CDD and Qhull are considered.

The C and C++ implementation of Double Description method (CDD) computes dual convex-hull, where the vertices of a polytope are defined as an intersection of half-spaces. This way, the program computes convex-hulls of a point set as well as unbounded intersections of half-spaces. It provides a post-processing step to determine the entire face structure of the polytope, and an auxiliary program to

projects of the polytope to lower-dimensional subspaces. The numerical precision is handled either using an exact or floating arithmetic option (Fukuda, 1996).

The Qhull is a C implementation that combines the two-dimensional Quick-hull algorithm with the multi-dimensional Beneath-Beyond method, the dual of the Double-Description method and similar to the randomized algorithms (Clarkson, 1995). The program also computes convex-hull and half-space intersections. It is efficient in handling high number of data with less computer memory. The efficiency is further demonstrated in the Table 3.1, which compares the cost of computing the convex-hull of random data matrices using GBT and Qhull. The computation is performed in a Personal Computer with Pentium IV processor and 256 MB RAM.

Table 3.1 Comparison of Computational Cost between GBT and Qhull

Matrix size	Average CPU seconds	
	GBT	Qhull
20 x 3	0.21	0.11
200 x 3	2.61	0.11
2000 x 3	47.84	0.17

Qhull is selected for implementation of the framework due to the availability of documentation's, technical support and compatibility with MS Windows platform (Barber, 1996), while CDD is for use within Unix platform (Fukuda, 1996). It is worth mentioning that very recently, the CDD interface to MATLAB in MS Windows has been made available in experimental version (Baotic and Torrisi, 2002), and may be considered within the proposed framework in future developments.

The Qhull interface function also has been available in MATLAB ver. 6.0, specifically the function *convhulln.m* (Mathworks, 2000b). However, this function returns the triangulated facets of the convex-hull, while the interest of the proposed framework is the vertices. The computation of intersection is also unavailable. Therefore, several

functions to facilitate geometric computations in MATLAB have been developed in this study, namely *chullne.m, qhalfne.m, qhalfne2.m, makedos.m, makeeds.m* and *disext.m.*

In providing geometric solution, these functions receive inputs in form of matrices. The functions translate the matrices into ANSI text files with suitable formats for the respective Qhull operations. The operations are invoked along with the specified options and data files, and the results are returned in form of text data files. The functions read and convert the data into matrices, then return the matrices to the caller.

The *chullne.m* computes the convex-hull of a matrix data set. Here, the input is a single matrix with size $n_o \times n_g$, representing n_g-dimensional data for n_o-observations. The function returns two outputs, the index of the convex-hull vertices **k**, and the convex-hull volume **v**. These are produced by invoking Qhull to compute the convex-hull vertices and volume with triangulation method. Therefore, this function can be used to determine AOS_θ convex-hull and volumes, as well as the GAE and GIAE.

The *qhalfne.m* computes the convex hull of an intersection of two data set matrices. Here, the inputs are two matrices with the same dimension, but may have different numbers of observations. The function returns two properties of the convex-hull of the intersections, which are the indices **k,** and the volume **v**. There are several steps back and forth between this function and Qhull. Firstly, the function sends the formatted text file of the matrices to Qhull, which returns the converted matrices in form of half-spaces. The function combines the half-spaces, nominates a common feasible point, and sends them back to Qhull. This time Qhull builds the intersection of half-spaces and returns the convex-hull indices and volume. The function converts them into matrices for further use within the framework. This can be used to determine the intersection of two multi-dimensional spaces with the same dimension.

The intersection of AOS_θ and DOS could be computed with *qhalfne.m.* However, the computational efforts especially within the outer level can be saved by noting that DOS is fixed throughout the framework. Therefore, the ANSI text file of DOS halfspaces is prepared earlier by *makedos.m.* Then *qhalfne2.m,* which is the modification

of *qhalfne.m,* computes the intersection using the matrix of AOS_θ and the half-space file of DOS as inputs.

The AOS_θ is constructed by the member of EDS. The combination in EDS is generated automatically with *makeeds.m*, which receives the range and intervals of each disturbance, then generate the unduplicated combinations. This function can be used for several disturbance variables with different ranges and discretisation intervals.

Within the inner-levels, all disturbance combinations, which are generated by *makeeds.m*, are used to generate AOS_θ. It is performed by multiple dynamic simulations for each combination. The dynamic profiles, along with the respective simulation time step and disturbance combination are recorded continuously into a matrix. Then disturbance extraction is performed by *distext.m*. Here, the part of the matrix, containing only dynamic profiles and time steps, is sent for convex-hull computation. The corresponding indices of the convex-hull are then used to identify the critical disturbance combination in the original matrix. This procedure can be used for multiple dynamic simulations, each with variable time steps.

These functions are organised in a toolbox, called DOF Toolbox, to facilitate the implementation of the proposed Dynamic Operability Framework in MATLAB. They are also in ongoing development for general public use.

3.4.3 Comparison with Parallel Studies

Recently, Ierapetritou (2001) also utilised Qhull to quantify process feasibility in terms of Feasible Convex-Hull Ratio (FCHR). The index is defined as the ratio between the volume of the feasible convex hull and volume of the overall expected range, similar to AOS_θ and DOS. The expected range is constructed in similar way to DOS. However, the feasible convex hull is constructed from the solution of the optimisation problem of finding maximum deviation due to the expected disturbances. The optimisation is the classic flexibility index problem (Grossmann *et al.*, 1983), and is applied to analyse the feasibility of steady state processes.

Meanwhile, Georgakis *et al.* (2001) expanded the steady state Output Controllability Index to Operability Index, covering dynamic servo and regulatory performance. DOS and AOS are expanded to the time horizon and are renamed the Dynamic Desired Operating Space (dDOpS) and Dynamic Achievable Operating Space (dAOpS), respectively. The dDOpS is the desired dynamic performance within the maximum allowable response time. The dAOpS is the best achievable dynamic performance for a given choice of inputs, desired outputs and expected disturbances using an idealised optimal controller. To measure operability index, two more spaces, S_1 and S_2, are introduced. The operating space S_1 is a combination of set points and the expected disturbance space. The achievable response space S_2 is a combination of set points and expected disturbances achieved within a maximum allowable response time. S_2 is obtained as a projection of intersection between dDOpS and dAOpS on to S_1. Then, Dynamic Operability Index dOI is defined as the ratio between the size of S_2 and S_1. In principle, the index assesses a particular process performance relative to the perfect control performance.

Both Ierapetritou (2001) and Georgakis *et al.* (2001) approached the flexibility and controllability problem differently in comparison to the proposed framework. Ierapetritou calculated the convex hull after the optimisation process was completed, while this framework utilises the convex-hull as one of the constraints during the optimisation process. Georgakis excluded the time dimension on S_1 and S_2, hence it was comparable to flexibility analysis within the proposed framework. The dOI is also based on the model invertibility and relative assessments against perfect control performance. In contrast, the proposed framework does not require model inversion and focuses on realisable control performance. Therefore, the current stage of the proposed framework maintains the independent ideas and approaches in relation to these parallel studies.

3.5 Case Study

3.5.1 Problem Formulation

This section presents the application of the proposed framework in a case study from Bahri (1996). It consists of two Continuous-Stirred-Tank-Reactors (CSTR) in series with an intermediate mixer introducing the second feed, as shown in Figure 3.4. The CSTRs are called reactors from this point onwards. An irreversible, exothermic and first order reaction of A→B takes place in both reactors. There are cooling jackets surrounding each reactor providing heat control for the reactions. The nonlinear dynamic model of the process is derived from material and energy balances around different parts of the process, assuming negligible mixer dynamics.

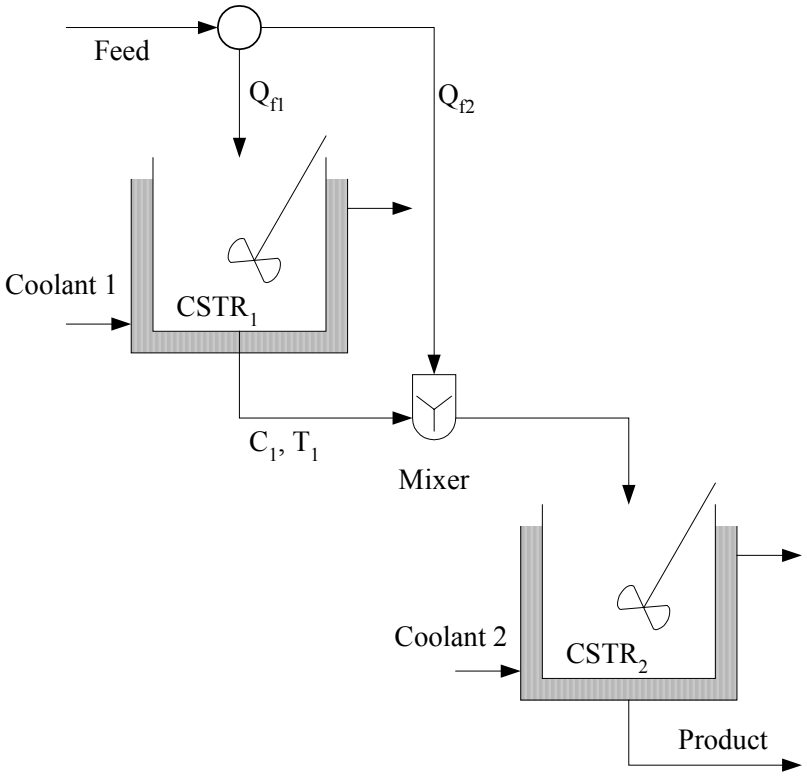


Figure 3.4 Case Study: Two Continuous Stirred Tank Reactors with a Mixer

The process model is as follows:

$$V_1 \frac{dC_1}{dt} = -k_o e^{-E/RT_1} C_1 V_1 + Q_{f1}(C_f - C_1) \tag{3.19}$$

$$Cool_1 = \frac{U_a m_{c1} cp}{U_a + m_{c1} cp}(T_{c1} - T_1) \tag{3.20}$$

$$V_1 \frac{dT_1}{dt} = D_h k_o e^{-E/RT_1} C_1 V_1 + Q_{f1}(T_f - T_1) + Cool_1 \tag{3.21}$$

$$C_m = \frac{(Q_{f1} C_1 + Q_{f2} C_f)}{Q_m} \tag{3.22}$$

$$T_m = \frac{(Q_{f1} T_1 + Q_{f2} T_f)}{Q_m} \tag{3.23}$$

$$Q_m = Q_{f1} + Q_{f2} \tag{3.24}$$

$$V_2 \frac{dC_2}{dt} = -k_o e^{-E/RT_2} C_2 V_2 + Q_m(C_m - C_2) \tag{3.25}$$

$$Cool_2 = \frac{U_a m_{c2} cp}{U_a + m_{c2} cp}(T_{c2} - T_2) \tag{3.26}$$

$$V_2 \frac{dT_2}{dt} = D_h k_o e^{-E/RT_2} C_2 V_2 + Q_m(T_m - T_2) + Cool_2 \tag{3.27}$$

The reaction depends on the feed flow into the first reactor Q_{f1}, and the feed flow into the mixer Q_{f2}. The measured variables are the temperatures T_1, T_2, the amount of heat extracted from each reactor $Cool_1$, $Cool_2$, and the product composition C_2. The heat transfer between the reactors and the jackets depends on the coolant flow rate m_{c1} and m_{c2}, coolant temperature T_{c1} and T_{c2}, coolant heat capacity cp and reactor heat transfer coefficient Ua, all with fixed values. The expected disturbances are the feed temperature T_f and the feed composition C_f. They are assumed as sets of step changes with uniformly distributed magnitudes. The lower bounds, nominal values and upper bounds of the disturbances are given in Table 3.2.

Table 3.2 The Expected Disturbance Combination

	Lower Bound	Nominal Value	Upper Bound
T_f (K)	298	T_{fN} = 300	315
C_f (mol/m^3)	19.5	C_{fN} = 20	21

The process is optimised to obtain an optimum nominal net profit $\Phi_{ss}=\Phi(t=t_0)$. The profile Φ is a function of design and dynamic variables as follows:

$$\Phi = 10(Q_{f1}C_{fN} + Q_{f2}C_{fN} - 0.3(Q_{f1} + Q_{f2})) - 0.01Cool_1 - 0.1Cool_2 \cdots - 0.1(Q_{f1} + Q_{f2}) \quad (3.28)$$

Here, C_{fN} is the nominal feed concentration, 20 mol/m^3.

The optimisation is subject to the following inequality constraints:

$$\begin{aligned} &g_1 : Q_{f1} \geq 0.05 && g_2 : Q_{f2} \geq 0.05 \\ &g_3 : Q_{f1} + Q_{f2} \leq 0.8 && g_4 : C_2 \leq 0.3 \\ &g_5 : T_1 \leq 350 && g_6 : T_2 \leq 350 \\ &g_7 : Cool_1 \leq 30 && g_8 : Cool_2 \leq 20 \end{aligned} \quad (3.29)$$

The process is analysed through an open-loop and two closed-loop cases. The first closed-loop case utilises only one Proportional–Integral (PI) controller to regulate T_1 by manipulating the coolant flowrate m_{c1}. The second closed-loop case utilises one more PI controller to regulate $Cool_2$ by manipulating the coolant flowrate m_{c2}. In the closed-loop cases, m_{c1} and m_{c2} are not fixed anymore. These variables oscillate around their steady state values m_{c1ss} and m_{c2ss}. The respective control equations are as follows:

$$e_1 = (T_{1ss} - T_1) \quad (3.30)$$

$$\frac{dI_1}{dt} = e_1 \quad (3.31)$$

$$m_{c1} = m_{c1ss} + \alpha_1 K_{c1}\left(e_1 + \frac{I_1}{\tau_1}\right) \quad (3.32)$$

$$e_2 = \left(Cool_{2ss} - Cool_2\right) \tag{3.33}$$

$$\frac{dI_2}{dt} = e_2 \tag{3.34}$$

$$m_{c2} = m_{c2ss} + \alpha_2 K_{c2}\left(e_2 + \frac{I_2}{\tau_2}\right) \tag{3.35}$$

Here, e_1, e_2 are the error signals, I_1, I_2 are the integral of the errors. All PI controllers are tuned using the Ziegler-Nichols method (Bahri, 1996; Ziegler and Nichols, 1943), and the parameters K_{C1}, K_{C2}, τ_1 and τ_2 are given in Table 3.3. Further tuning is facilitated using the scaling factors α_1 and α_2.

Table 3.3 Controller Parameters

Control loop		Proportional gain K_c	Reset times τ
No	Variable pair		
1	T_1 - m_{c1}	0.0023	0.1667
2	$Cool_2$ - m_{c2}.	0.01456	0.1668

Therefore, there are three sets of process models. All cases contain the open-loop DAE (3.19)-(3.27). In addition, the single PI controller case contains equations (3.30)-(3.32), and the double PI controllers case contains equations (3.33)-(3.35). Considering the control purpose as well as optimisation objective and constraints, $Cool_1$, $Cool_2$, m_{c1}, m_{c2} and net profit (3.28) become the outputs of DAE, in addition to the state variables T_1, T_2, C_1, C_2, I_1 and I_2.

The framework therefore seeks out the optimum values of Q_{f1}, Q_{f2}, α_1, α_2, as well as the initial values T_{1ss}, T_{2ss}, C_{1ss}, C_{2ss}, $Cool_{1ss}$, and $Cool_{2ss}$. The initial values of coolant flowrates m_{c1ss} and m_{c2ss} are fixed, while the initial values of integral of errors I_{1ss} and I_{2ss} are naturally zero. Therefore, these are not considered as design variables.

3.5.2 Framework Implementation

Within the proposed framework, the models, objective and constraints are arranged according to the formulations (3.13)-(3.14) and (3.17)-(3.18). All outer-levels contain both steady state and DAE models. The steady state model determines the process initial conditions at the nominal disturbance values, and therefore, does not include any control actions. Hence, the steady state version of (3.19)-(3.28) is used. It is followed by DAE for the respective cases.

The inequality constraints $\mathbf{g_{1\text{-}3}}$ determine the design variable values, which are once decided, are fixed during the dynamic simulation. Therefore, $\mathbf{g_{1\text{-}3}}$ belong to $\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le 0, \mathbf{j} \in \mathbf{I_r}$ in (3.13). The constraints $\mathbf{g_{4\text{-}8}}$ apply to both the initial conditions and the dynamic profiles of T_1, T_2, $Cool_1$, $Cool_2$ and C_2. The initial conditions T_{1ss}, T_{2ss}, $Cool_{1ss}$, $Cool_{2ss}$ and C_{2ss} are considered as design variables and constrained by $\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le 0, \mathbf{j} \in \mathbf{I_r}$, while $T_1(t)$, $T_2(t)$, $Cool_1(t)$, $Cool_2(t)$ and $C_2(t)$ over the time horizon t are constrained by DOS. Therefore, the optimisation constraints are rearranged as follows:

$$\mathbf{g_{1-8}} = \begin{cases} g_1 : Q_{f1} \ge 0.05 & g_2 : Q_{f2} \ge 0.05 \\ g_3 : Q_{f1} + Q_{f2} \le 0.8 & g_4 : C_{2ss} \le 0.3 \\ g_5 : T_{1ss} \le 350 & g_6 : T_{2ss} \le 350 \\ g_7 : Cool_{1ss} \le 30 & g_8 : Cool_{2ss} \le 20 \end{cases} \tag{3.36}$$

$$\mathbf{DOS} = \begin{cases} 0 \le C_2(t) \le 0.3 \\ 0 \le T_1(t) \le 350 \\ 0 \le T_2(t) \le 350 \\ 0 \le Cool_1(t) \le 30 \\ 0 \le Cool_2(t) \le 20 \\ t_0 \le t \le t_f \end{cases} \tag{3.37}$$

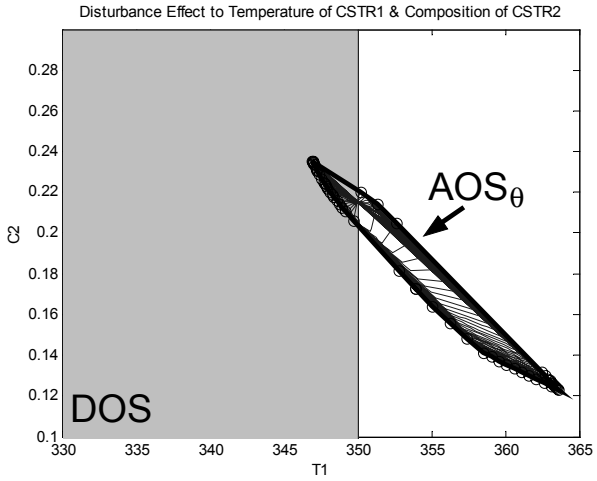
The AOS_{θ} is constructed from the six variables defining DOS. The process dynamics are simulated over 100 seconds time horizon with 1-second intervals.

3.5.3 Results and Discussion

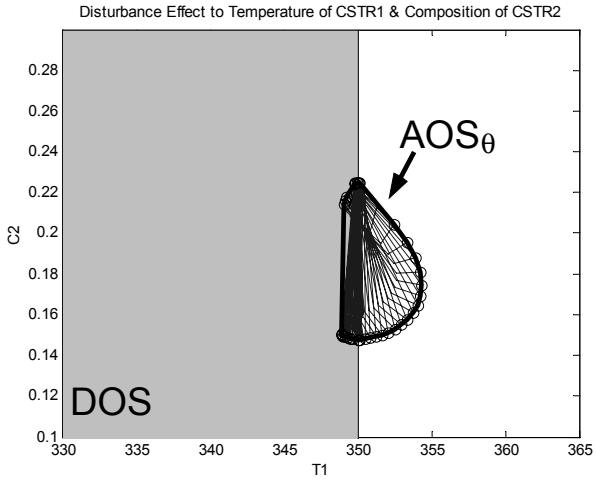
The proposed framework is applied to the case study and the results are compared to that from the original framework. Similar to the original, the first outer-level in the proposed framework supplied optimal open-loop operating conditions. In the first inner-level, disturbances are applied to the process and controllers are used where applicable. The AOS_θ formed in this level is projected back to EDS to deliver all disturbance combinations causing the maximum deviations relative to the operating point.

Since the process is non-convex, it is interesting to see how the points inside EDS contribute to AOS_θ. For this purpose, EDS is discretised into equally spaced grids, and AOS_θ is constructed by projecting all the grid points into the output space. Since visualisation of 6-dimensional space is impossible, the dynamics of AOS_θ are initially observed through two 2-dimensional variable spaces C_2-T_1 and $Cool_2$-T_2 at every time intervals. These spaces are selected because the original approach had shown that constraint violations occurred on those during the iterations. Figures 3.5-3.6 show how the 2-dimensional convex-hulls of AOS_θ evolved for all cases in the first inner-level. The convex-hull evolved from an initial operating point to a nearly rectangular curve that expanded and finally settles in open-loop case, or shrink before settling in closed-loop case. They also rotate due to interactions between variables. The thick outer line in the figures is the projected convex-hull of AOS_θ for the whole time horizon into the respective spaces.

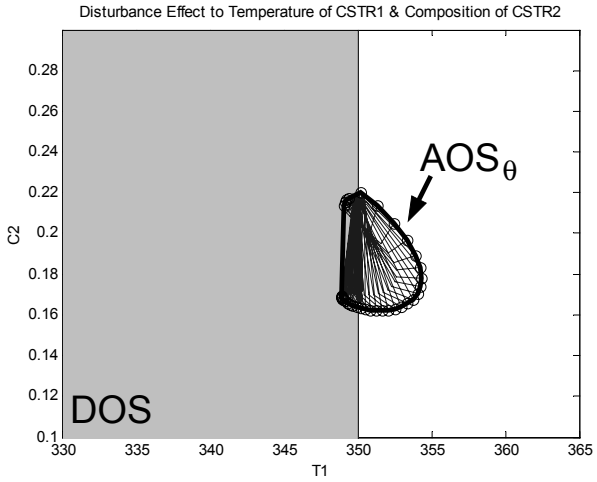
The extraction of critical disturbances of the open-loop based on the 6-dimensional AOS_θ, however, are not instantly successful. The AOS_θ constructions of open-loop case is initially failed because Qhull detected that respective AOS_θ had less dimensions and therefore demanded manual elimination of several variables. The same construction at the single PI-control and double-PI control went through, but returned incorrect disturbance combinations, which include those inside the boundaries.



(a)



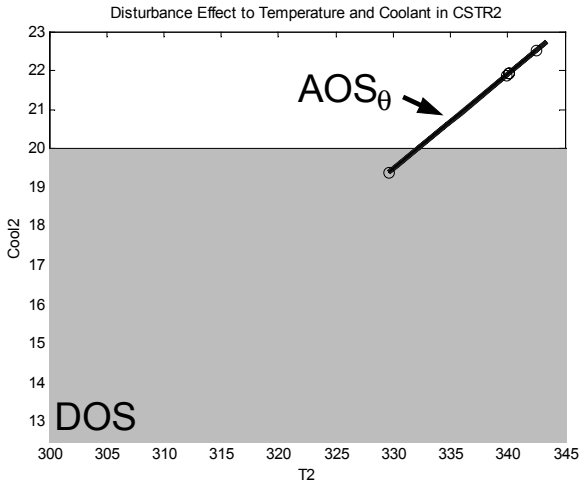
(b)



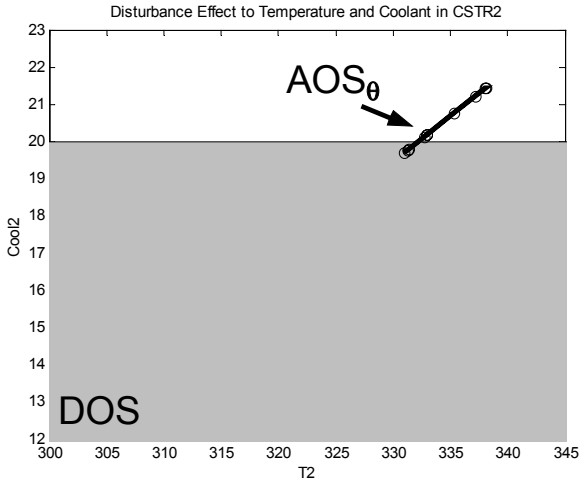
(c)

Figure 3.5 The Projection of AOS_θ into $C_2 - T_1$ Space after Iteration 1

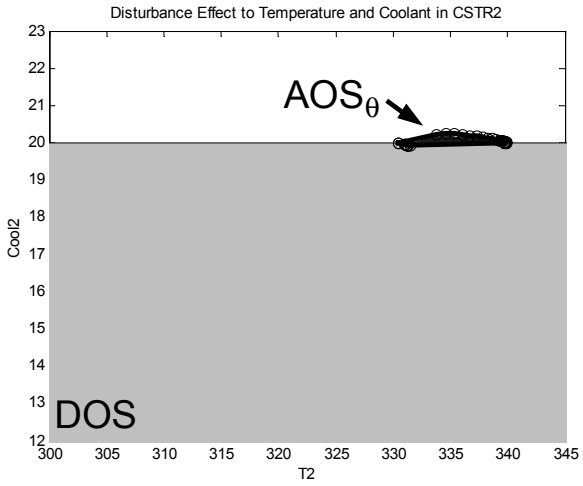
(a) Open-Loop (b) Single-PI Control (c) Double-PI Controls



(a)



(b)



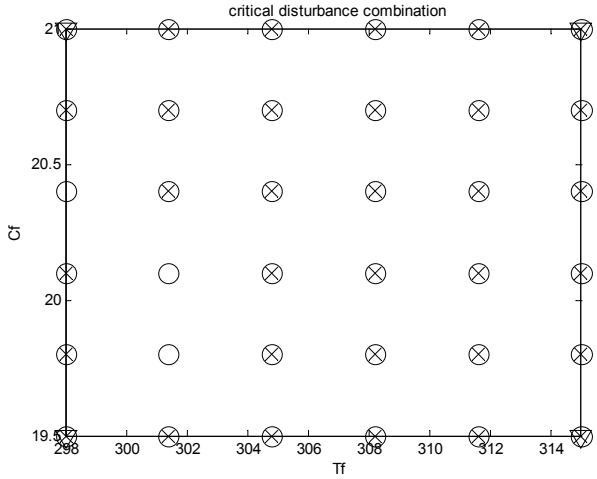
(c)

Figure 3.6 The Projection of AOS_θ into $Cool_2 - T_2$ Space after Iteration 1

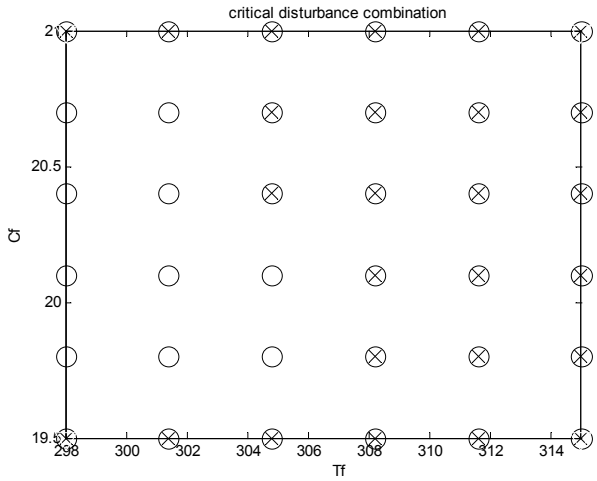
(a) Open-Loop (b) Single-PI Control (c) Double-PI Controls

The Critical Disturbance Extraction
is Failed

(a.i)



(b.i)



(c.i)

Figure 3.7 The Critical Disturbance Found in the First Inner-Level

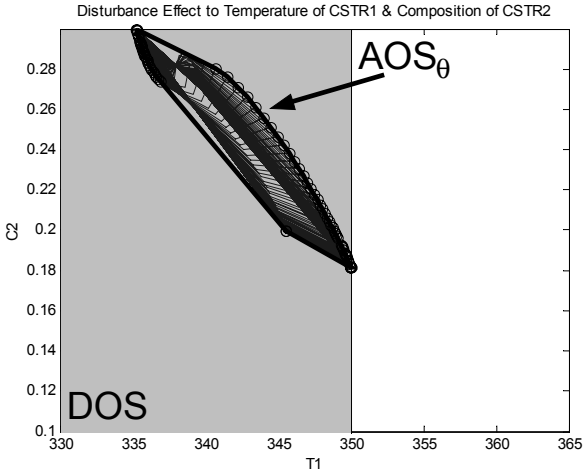
(a) Open-loop (b) Single-PI control (c) Double-PI Controls

◦ = EDS Member × = Detected Combination ∇ = Critical Combination

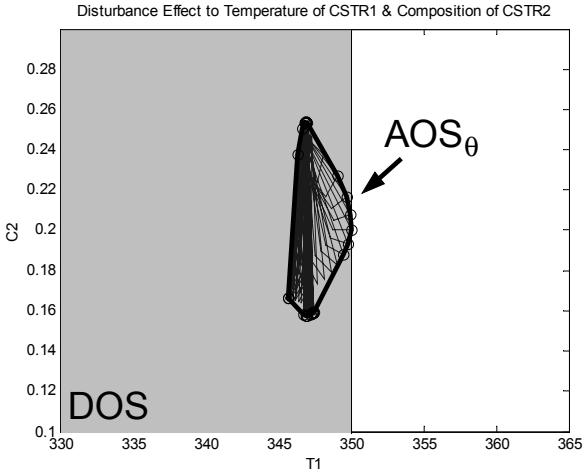
The suspected variables are the pairs $Cool_1$-T_1 and $Cool_2$-T_2 for the open-loop case, and $Cool_2$-T_2 for the single-PI control case. The problem is due to linear relationships (collinearity) within each pair, as illustrated in Figures 3.6a-b for $Cool_2$-T_2 and inferred accordingly for $Cool_1$-T_1. Those figures show that the respective planes are reduced into line segments (coplanar), which means only one dimension is required to represent both variables. Apparently, control actions break the linearity between the pair, as shown in Figure 3.6c, and this is supported by the fact that the dimensional problem does not occur on the double-PI control case. This indicates that AOS_θ should be constructed by fewer variables when both Cool and T in one reactor are not controlled. Hence, one of them is redundant, and shall be eliminated.

To provide comparable conditions, the elimination of redundant variables is considered to suit all cases. Further examination of Figure 3.6 shows that $Cool_1$ and T_2 are always within the feasible region, even when their counterparts T_1 and $Cool_2$ are not controlled. This fact is also supported by analytical evaluation of the steady state model against the respective variable constraints. Therefore, these output variables are eliminated from AOS_θ in all cases. This manual elimination of redundant variables is nonetheless favorable within the proposed framework, therefore an automatic elimination is considered for further development.

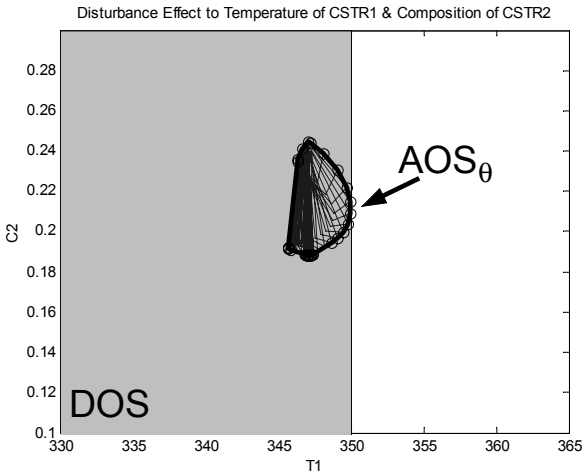
The remaining variables T_1, $Cool_2$, and C_2 successfully construct the AOS_θ, leading to the projection back to EDS. The outer-level in the second iteration considers these disturbance combinations and solves the optimisation problem. Figures 3.8-3.9 show that the projected AOS_θ for all cases after the second iteration are located completely inside DOS. The subsequent inner-levels do not find new combinations, therefore all cases are completed in two iterations.



(a)



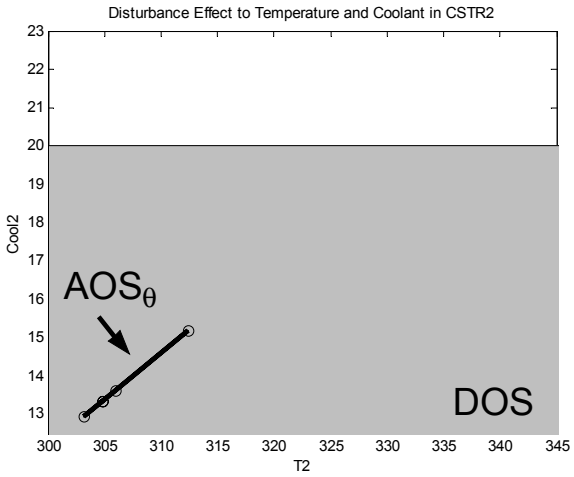
(b)



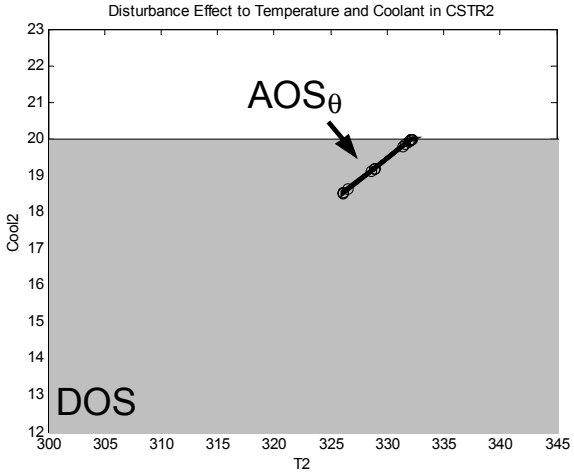
(c)

Figure 3.8 The Projection of AOS_θ into $C_2 - T_1$ Space after Iteration 2

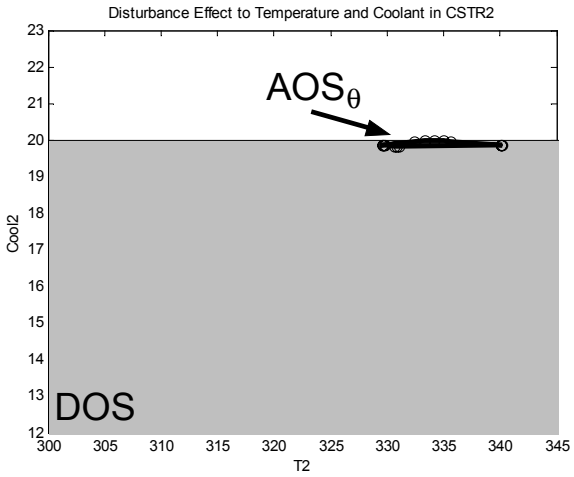
(a) Open-loop (b) Single-PI Control (c) Double-PI Controls



(a)



(b)



(c)

Figure 3.9 The Projection of AOS_θ into $Cool_2 - T_2$ Space after Iteration 2

(a) Open-Loop (b) Single-PI Control (c) Double-PI Controls

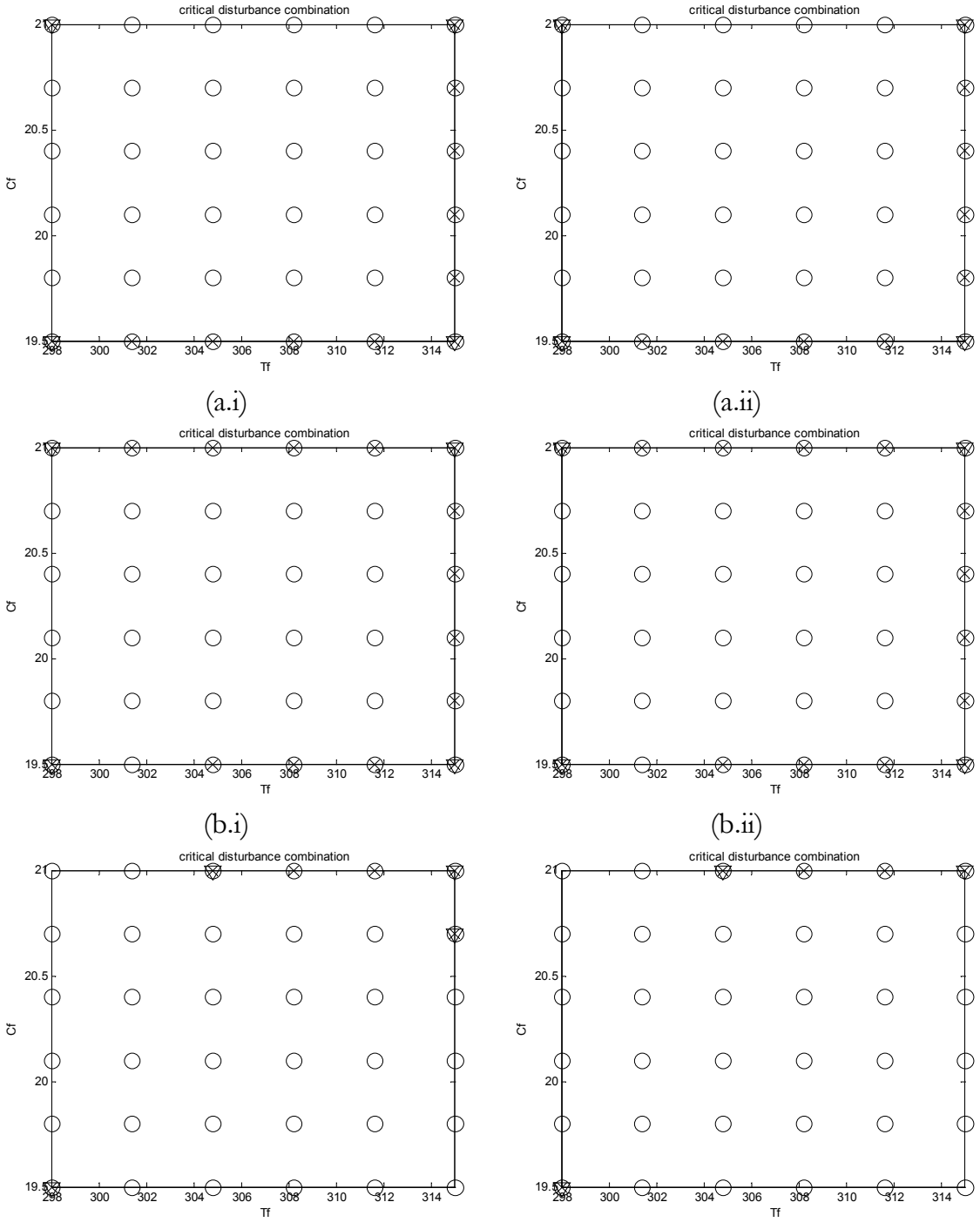


Figure 3.10 The Critical Disturbance Found in Both Inner-Levels

(a) Open-Loop (b) Single-PI Control (c) Double-PI Controls
(i) Iteration 1 (ii) Iteration 2
∘ = EDS Member × = Detected Combination ∇ = Critical Combination

Figure 3.10 shows the comparison between the disturbances extracted in the first and second inner levels. All extracted critical disturbances using the selected variables are located at the boundaries of disturbance space, which are logically expected because all of these cases are monotonic in EDS. However, those are not specifically located in the vertices, in comparison with earlier statement by Halemane and Grossman (1981a,c in Grossmann *et al.*, 1983) as stated in section 3.3.2. The problem is suspected due to further redundancy in the selected variables, since the selection is based on the known constraint violations, which coincident with the collinear variables. This requires further analysis and a mechanism of redundancy elimination, which are addressed in the next chapter. At this stage, the critical disturbances are further extracted from the vertices of the detected combination.

Table 3.4 shows the extracted critical disturbances, while those found by the original approach are in boldface. The combination found by the proposed approach represents the largest deviation from the initial conditions, which include the variations within the feasible space. Therefore, there are more combinations than those found by the original approach.

Table 3.4 The Disturbance Combinations

	Iteration 1 $[T_f, C_f]$	Iteration 2 $[T_f, C_f]$
Dynamic open-loop	**[315, 21],** [315, 19.5], [298,19.5], [315,19.5]	**[315, 21],** [315, 19.5], **[298,19.5],** [315,19.5]
Dynamic closed-loop, single-PI control	**[315, 21],** [298,19.5], [298,19.5], [315,19.5]	**[315, 21],** [298,19.5], [298,19.5], [315,19.5]
Dynamic closed-loop, double-PI controls	**[315, 21],** [298,19.5], [304.8,21], [315,20.7]	**[315, 21],** [298,19.5] [304.8,21],

It is also worth noting the difference between the collinearity in the open-loop and the single-PI control cases. At the open-loop case, all collinear variables $Cool_1$-T_1 and $Cool_2$-T_2 are monotonic relative to all other variables, therefore the lower dimensional (degeneracy) problem is severe and easily detected. The collinearity in single-PI control case, on the other hand, is only between $Cool_2$-T_2. Furthermore, the respective profiles are not monotonic, since both variables oscillate due to the controlled T_1. The degeneracy problem is less obvious to Qhull, therefore the convex-hull computation proceeds. However, this time the resulting AOS_θ convex-hull consists of many coplanar data points in $Cool_2$-T_2 plane. These are projected back to almost all members of EDS, yielding incorrect extraction of critical disturbances and uncertainties.

The computational costs are 173.89 CPU seconds for the open-loop, and 4596.86 CPU seconds (12 minutes) for the closed-loop cases. These costs are not significantly different to the original approach, due to the number of critical disturbance combinations considered in the second outer levels. Nevertheless, the framework covers the important disturbance combinations in the second level, hence the solution is obtained within two iterations. Eliminating redundant variables may reduce the number of extracted critical disturbances and the associated computation time, and this procedure is addressed in the next chapter. At this development stage, the optimisation results are given in Table 3.5.

Table 3.5 The Controllability Assessment Results

	Q_{fi}	Q_{f2}	Φ_{ss} ($/hr)	GAE	GIAE	PVAR ($)
Steady state optimum (iter#1)	0.355	0.206	90.35			
Dynamic open-loop (iter#2)	0.252	0.055	46.86	0.0592	2.9021	189.515
Dynamic closed-loop, single-PI control (iter#2)	0.324	0.175	79.29	0.0478	2.7057	147.725
Dynamic closed-loop, double-PI controls (iter#2)	0.324	0.223	87.77	0.0116	0.4932	15.032

iter# = iteration number

The proposed framework maintains similar solutions to the original framework in terms of Q_{f1}, Q_{f2} and $\mathbf{\Phi}_{ss}$ for all cases. Several additional measurements are also provided, which are the approximate value of General Absolute Error (GAE) and General Integral Absolute Error (GIAE), and the Profit Variation (PVAR). They are computed from the recorded process dynamics in the inner-levels. Note that the inner-levels dynamics are resulted from the initial conditions provided by the preceding outer-levels. Therefore, the measurements represent the optimum dynamics quantity of the respective iterations.

The GAE is approximated by the volume of AOS_{θ}, which is computed by excluding the time horizon t, hence captures the process flexibility. The GIAE, in contrast, includes the time horizon, hence captures the controllability. The effect of process dynamics on profit profiles is quantified by the convex-hull of all profit profiles, which approximates the following accumulated variations over time and over disturbance combinations:

$$\mathrm{PVAR} = \iint_{t,\theta} \Phi\left(\mathrm{Cool}_1(t,\theta), \mathrm{Cool}_2(t,\theta)\right) d\theta dt \tag{3.38}$$

Therefore, GAE and GIAE complement the r-OCI to quantify the variable dynamics, while PVAR complements the nominal profit $\mathbf{\Phi}_{ss}$ to translate the dynamics into process economics.

The indices GAE, GIAE and PVAR are expectedly higher for dynamic open-loop cases, indicating higher sensitivity to disturbances. The respective nominal profit $\mathbf{\Phi}_{ss}$ is the smallest, since most of the resources are utilised to absorb the disturbance effects. The double-PI control strategy, as expected, provided the closest $\mathbf{\Phi}_{ss}$ to the nominal steady state and the smallest of all other indices. Interestingly, the single-PI control strategy does not perform better than the open-loop case in terms of GAE, GIAE and PVAR, although the change in $\mathbf{\Phi}_{ss}$ is significant.

This can be explained by observing the dynamics of the pair T_1-$Cool_1$ and T_2-$Cool_2$ subject to all disturbance combination, taken from the second inner-level of single PI-control case in Figures 3.11-3.12.

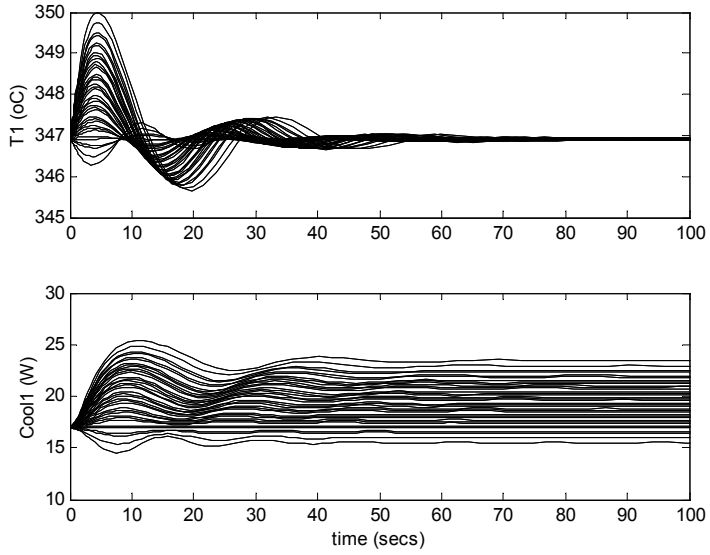


Figure 3.11 The Interaction between the Controlled T_1 with Uncontrolled $Cool_1$ Single-PI Control

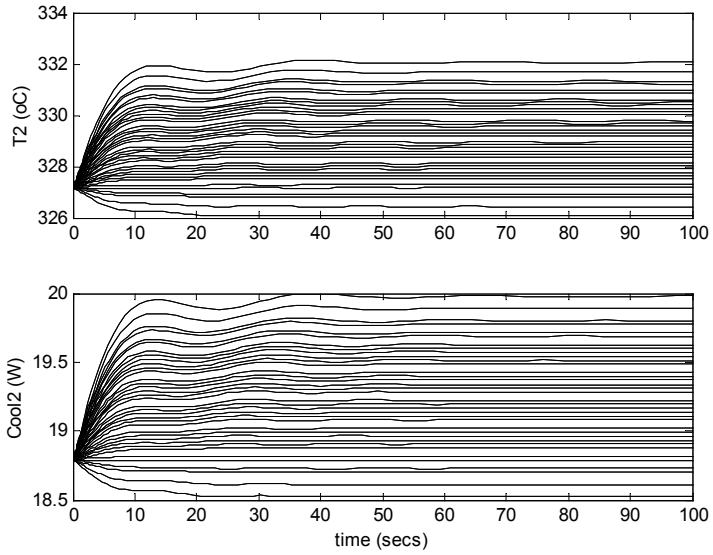


Figure 3.12 The Interaction between the Controlled $Cool_2$ with Uncontrolled T_2 Single-PI Control

Both figures show that controlling T_1 leaves offsets on the $Cool_1$. The oscillation of $Cool_1$ is acceptable, since the constraint on $Cool_1$ is relaxed in comparison to $Cool_2$ ($Cool_1 \leq 30$, $Cool_2 \leq 20$). Therefore, T_1 can be pushed significantly close towards its constraint without causing the infeasibility in $Cool_1$. This significantly increases the values of Qf_1, T_{1ss}, $Cool_{1ss}$, and eventually higher $\mathbf{\Phi}_{ss}$, in comparison to the open-loop case.

However, the values of GAE, GIA and PVAR do not accordingly reduced. The GAE and GIAE calculations involve T_1, $Cool_2$ and C_2. The value of T1 is dominating other variables. Since in single-PI control the second reactor is left uncontrolled, there are significant variations in $Cool_2$ and C_2, as shown in Figures 3.8b-3.9b. The combination of these variations with the dominant T1 yields significantly large GAE and GIAE. Meanwhile, PVAR calculation involves $Cool_1$ and $Cool_2$. Again the high offsets in $Cool_1$ and $Cool_2$ as shown in Figures 3.11-3.12 explain why PVAR of single-PI control is large and significantly close to the open-loop case.

Controlling $Cool_2$ in the double-PI control case significantly reduces the variation in $Cool_2$ in comparison to the other two cases, as shown in Figures 3.13 and 3.9c. This causes the reduced variations in C_2, as indicated in Figure 3.8c. Therefore, the combined dynamics of T_1, $Cool_2$ and C_2 yields the lowest GAE and GIAE. Accordingly, PVAR is also significantly reduced, although still persistent due the oscillations and offsets on $Cool_1$.

The comparison of the values of GAE, GIAE and PVAR between the three control cases highlights the different quality of process dynamics within the DOS. Note that r-OCI within the proposed framework guarantees the complete process flexibility and controllability. Therefore, the variation occurs completely within the feasible operating region, where the indices rank the relative significance of the control actions to process dynamics and economy. This emphasises the complementary roles of GAE and GIAE to r-OCI, as well as the profit variation PVAR to the flexibility cost $\mathbf{\Phi}_{ss}$.

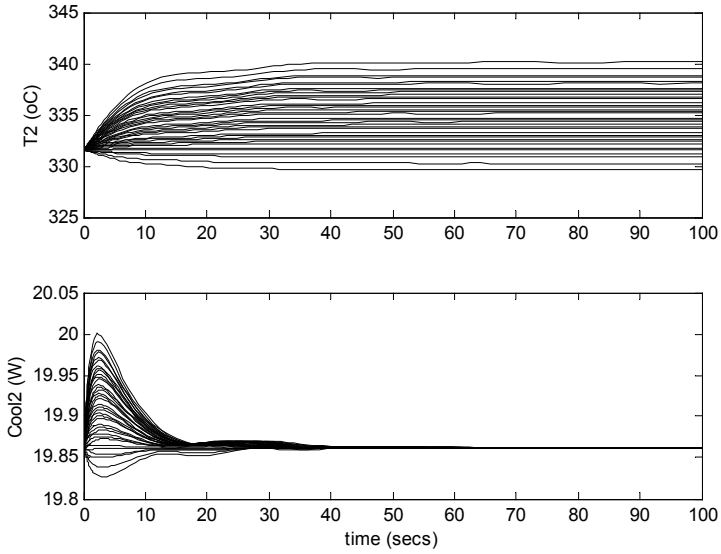


Figure 3.13 The Interaction between the Controlled $Cool_2$ with Uncontrolled T_2 Double-PI Controls

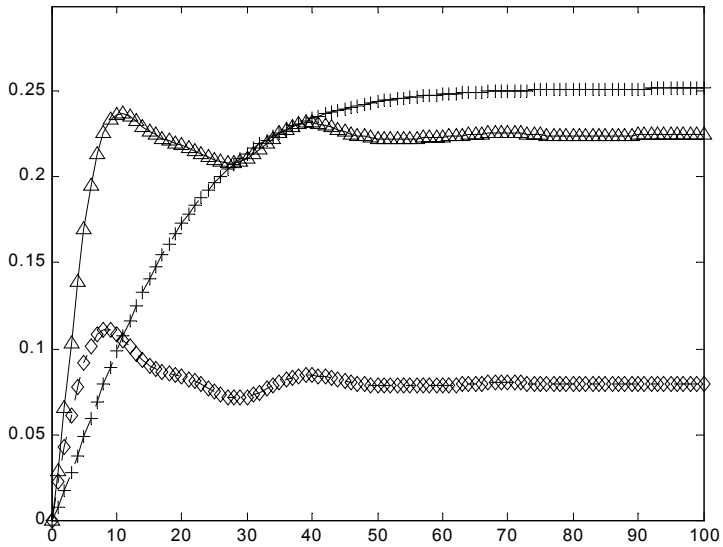


Figure 3.14 The Profit Variation over Time Horizon

+: Open-Loop, Δ: Single-PI Control , ◊: Double-PI Controls

The three control cases in this study suggest extending the problem to control structure selection, where the control pairs may be selected based on economic or dynamic performance objectives. For this particular process, controlling T_1 is mandatory to satisfy both objectives. It is supported by an experiment of controlling $Cool_1$ using m_{c1}. Both the original and the proposed framework are converged with the controller scaling factor α_1 equal to zero, meaning it is preferable to leave this loop open. It would be interesting to see whether the result would be reproduced during an automatic structure selection. This is considered for the next stage of the framework development, involving the application of the Mixed Integer Nonlinear Programming (MINLP).

3.6 Conclusion

An integration of the Output Controllability Index within the Dynamic Operability Framework (DOF) for regulatory cases of a fixed process and controller structure has been presented. The approach proposes the utilisation of a geometric representation of a feasible operating region, and a projection of disturbance space into output space to assess the controllability. Within the proposed approach, flexibility becomes the special case of the controllability problem. The approach replaces the multiple maximisation problems in the inner-level with a set of geometric operations. Several indices taking account of the interactions between dynamic variables as well as translation to process economics are provided, which are the General Absolute Error (GAE), General Integral Absolute Error (GIAE), and Profit Variation (PVAR). The case study has demonstrated the proposed framework capability to assess dynamic and nonlinear chemical processes.

The framework implementation involves the selection of dynamic solver and optimiser, as well as the development of several DOF Toolbox functions to facilitate the geometric and indices computation. This toolbox is expanding along with the further development of the proposed framework.

Several issues have been noted during the current stage. These are the development of high-index DAE solvers and global optimisers, the elimination of redundant dynamic variables, the selection of process and control structures selections and further analysis of critical disturbance combinations. The solver and optimiser issues are highly specialised both mathematically and algorithmically, therefore they are considered as separate research issues beyond the coverage of this thesis. The other issues are addressed in further stages of framework development, and presented in the next chapters.

3.7 Nomenclature

Acronyms		
AIS	:	Available Input Space
AOS	:	Achievable Output Space
AOS_θ	:	Achievable Output Space due to disturbances
BFGS	:	Achievable Output Space due to disturbances
CSTR	:	Continuous Stirred Tank Reactor
DAE	:	Differential Algebraic Equations
DIS_θ	:	Desired Input Space due to disturbances
DIS_w	:	Desired Input Space due to outputs
DOF	:	Dynamic Operability Framework
DOS	:	Desired Output Space
dDOpS	:	Dynamic Desired Operating Space
dAOpS	:	Dynamic Achievable Operating Space
dOI	:	Dynamic Operability Index
EDS	:	Expected Disturbance Space
FCHR	:	Feasible Convex-Hull Ratio
GAE	:	Generalised Absolute Error
GBT	:	Geometric Bounding Toolbox
GIAE	:	Generalised Integral Absolute Error
GMC	:	Generic Model Control
I/O	:	Input Output
IAE	:	Integral Absolute Error
ISE	:	Integral Squared Error
ITAE	:	Integral Time Absolute Error

Acronyms		
LP	:	Linear Programming
LQG	:	Linear Quadratic Gaussian
MIMO	:	Multi-Input Multi-Output
NLP	:	Nonlinear Programming
OCI	:	Output Controllability Index
s-OCI	:	Servo Output Controllability Index
r-OCI	:	Regulatory Output Controllability Index
ODE	:	Ordinary Differential Equations
PI	:	Proportional - Integral
PVAR	:	Profit Variation
PCG	:	Preconditioned Conjugate Gradient
QP	:	Quadratic Programming
SQP	:	Sequential Quadratic Programming

Variables		
C	:	Concentration of A in CSTR (mol/m^3)
Cool	:	Amount of heat removed from CSTR by coolant (KJ)
c_p	:	Heat capacity (J/Kg K)
Dh	:	Heat of reaction (5 K. m^3/mol)
E	:	Activation energy (6000 K)
e_1, e_2	:	Errors
$\mathbf{f_i}$	:	Set of ordinary differential equations
$\mathbf{g_j}$	:	Inequality constraints
$\mathbf{h_i}$	:	Equality constraints
I_1, I_2	:	Integral of errors
K	:	Proportional gain of PI controller
k_o	:	Reaction constant
M	:	Mass matrix
m_c	:	Coolant flow rate (m_{c1} = 0.35 m^3/sec, m_{c2} = 0.8 m^3/sec)
p	:	Process parameters
Q	:	Liquid Flow rate (m^3/sec)
R	:	Reaction rate (moles/sec)
T	:	Temperature of the reactor (K)
T_c	:	Temperature of the coolant (250 K)
t	:	Time (sec)
U	:	Set of manipulated variable values
UA	:	Overall heat transfer coefficient (0.35 W/K)
u	:	Vector of manipulated variables
V	:	CSTR volume (m^3)
W	:	Set of output/measured variable values
$\overline{\mathbf{W}}$	:	Augmented set of measured variable values
w	:	Vector of output/measured variables

Variables		
$\overline{\mathbf{w}}$	:	Vector of augmented measured variables
$\mathbf{X}$	:	Set of state variable values
$\mathbf{x}$	:	Vector of state variables
$\mathbf{Z}$	:	Set of design variable values
$\overline{\mathbf{Z}}$	:	Augmented set of design variable values
$\mathbf{z}$	:	Vector of design variables
$\overline{z}$	:	Vector of design variables

Greek Letters		
α	:	Scaling values for controller parameters
Φ	:	Objective function
Θ	:	Set of all disturbance realisations
μ	:	Volume of a multi dimensional space
θ	:	Vector of disturbances and uncertainties
τ	:	Integral/reset time of PI controller

Superscripts		
k	:	Critical values
l	:	Lower bounds
u	:	Upper bounds
N	:	Nominal values
*	:	Optimal solution from the last iteration

Subscripts		
ss	:	Initial conditions/ steady states
1	:	First CSTR
2	:	Second CSTR
d	:	Dynamic
f	:	Feed
m	:	mixer output

3.8 References

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Bahri, P. A., Bandoni, A. and Romagnoli, J. (1996). "Operability Assessment in Chemical Plants", Computers and Chemical Engineering, 20, S787 - S792.

Baotic, M. and Torrisi, F. (2002). "CDD-MATLAB Interface".

Barber, C. B. (1996). "Qhull", Minneapolis, USA, The Geometry Center, University of Minnesota.

Barber, C. B., Dobkin, D. P. and Huhdanpaa, H. T. (1996). "The Quickhull Algorithm for Convex-Hulls", ACM Transaction on Mathematical Software, 22(4), 469 - 483.

Berg, M. d., Kreveld, M. v., Overmars, M. and Schwarzkopf, O. (1997). "Computational Geometry, Algorithm and Applications", Sprieger.

Berzins, M., Dew, P. and Furzeland, R. (1989). "Developing Software for Time-Dependent Problems Using the Method of Lines and Differential-Algebraic Integrators", Applied Numerical mathematics, 5, 375 - 397.

Brenan, K., Campbell, S. and Petzold, L. (1996). "Numerical Solution of Initial-Value Problem in Differential-Algebraic Equations", Philadelphia, Society for Industrial and Applied Mathematics.

Cervantes, A. and Biegler, L. T. (1997). "Large-Scale DAE Optimization Using Simultaneous Nonlinear Programming Formulations", AIChE Journal, 44, 1038 - 1050.

Clarkson, K. L. (1995). "Hull".

Coleman, T. F. and Li, Y. (1996). "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds", SIAM Journal on Optimization, 6, 418-445.

Ekawati, E. and Bahri, P. A. (2001). "Adaptation of Output Controllability Index within Dynamic Operability Framework", 6th IFAC Symposium on Dynamics

and Control of Process Systems (DYCOPS 2001), Jejudo Island, Korea, Elsevier Science Ltd.

Ekawati, E. and Bahri, P. A. (2003). "Integration of the Output Controllability Index within the Dynamic Operability Framework in Process System Design", Journal of Process Control, 13(8), 717 - 727.

Fabian, G., van Beek, D. A. and Rooda, J. E. (2001). "Index Reduction and Discontinuity Handling Using Substitute Equations", Mathematical and Computer Modelling of Dynamical Systems, 7(2), 173 - 187.

Fukuda, K. (1996). "CDD", Zurich, Switzerland, ETH-Zentrum.

Georgakis, C., Uzturk, D., Subramanian, S. and Vinson, D. R. (2001). "On the Operability of Continuous Processes", 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-6), Chejudo Island, Korea.

Grossmann, I. E., Halemane, K. P. and Swaney, R. E. (1983). "Optimization Strategies for Flexible Chemical Processes", Computers and Chemical Engineering, 7(4), 439 - 462.

Hindmarsh, A. (1980). "LSODE and LSODI, Two New Initial Value Ordinary Differential Equation Solvers", ACM SIGNUM Newsletter, 10 - 11.

Ierapetritou, M. G. (2001). "New Approach for Quantifying Process Feasibility: Convex and 1-D Quasi-Convex Regions", AIChE Journal, 47(6), 1407 - 1417.

Logsdon, J. S. and Biegler, L. T. (1989). "Accurate Solution of Differential - Algebraic Optimization Problems", Industrial and Engineering Chemistry Research, 28, 1628 - 1639.

Mathworks (2000a). "Optimisation Toolbox User's Guide Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Mathworks (2000b). "Using MATLAB Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Shamos, M. (1978). "Computational Geometry", PhD Thesis, Yale University.

Shampine, L. F. and Reichelt, M. W. (1997). "The MATLAB Ode Suite", SIAM Journal on Scientific Computing, 18, 1 - 22.

Shampine, L. F., Reichelt, M. W. and Kierzenka, J. A. (1997). "Solving Index-1 DAE in MATLAB and Simulink", ACM Transaction on Mathematical Software.

Tjoa, I. B. and Biegler, L. T. (1991). "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential - Algebraic Equation Systems", Industrial and Engineering Chemistry Research, 30, 376 - 385.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Veres, S. M. (1999). "Iterative Identification and Control Redesign Via Unfalsified Sets of Models: A Basic Scheme", International Journal of Control, 72(10), 887 - 903.

Veres, S. M., Kuntsevitch, A. V., Hermsmeyer, S. and Wall, D. S. (1995). "Using GBT Version 5.1 in Identification and Control", 1995 MATLAB Conference.

Veres, S. M. and Mayne, D. (2001). "Geometric Bounding Toolbox", Southampton, UK, SysBrain.

Veres, S. M., Messaoud, H. and Norton, J. P. (1999). "Limited - Complexity Model-Unfalsifying Adaptive Tracking Control", International Journal of Control, 72(15), 1417 - 1426.

Vinson, D. R. and Georgakis, C. (1998). "A New Measure of Process Output Controllability", 5th IFAC Symposium on Dynamics and Control of Process Systems, Corfu, Greece.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10(2), 185 - 194.

Ziegler, J. G. and Nichols, N. B. (1943). "Process Lags in Automatic Control Circuits", Transactions of ASME, 65, 433.

Chapter 4

Redundancy Analysis and Elimination

4.1 Introduction

This chapter presents the solution of the redundancy problem of the proposed Dynamic Operability Framework encountered in previous chapter. The previous chapter presented the adaptation of the Output Controllability Index (Vinson and Georgakis, 2000), and the integration into the Dynamic Operability Framework (Bahri, 1996). The index computation involves the geometric operations between several high dimensional polyhedra associated with the measured dynamic output profiles. Some of these variables, however, may be collinear with each other. This collinearity relates to redundant dimensions in variable space, which disrupts the geometric computation, and causes incorrect detection of critical disturbance and uncertainty combinations within the framework.

This problem calls for a mechanism to eliminate the variable redundancies in a given data set. For that purpose, this chapter explores several methods to either address collinearity among variables, or reduce the dimensionality of a data set. Those methods are the Cross-Correlation Analysis, the Principal Component Analysis (PCA) and the Stepwise Collinear Diagnosis (SCD). Based on the features of these methods, an

algorithm for the redundancy analysis and elimination is developed. The algorithm groups the redundant variables, and selects one representative from each group. Several grouping and selection strategies are tested, to come out with the proper functional variables and correct critical disturbance combinations.

This chapter has been summarised along with the next chapter about the systematic process and controller structure selection, into a short conference paper. The paper is titled "Variable Redundancy Elimination and Automatic Structure Selection within Dynamic Operability Framework", and has been accepted at the Process System Engineering Symposium in Kunming, P. R. China, January 5th-9th 2004.

The structure of this chapter is as follows: Section 4.2 describes the problems leading to the requirement of redundancy elimination within the Dynamic Operability Framework. The reviews and adaptation of prospective methods follow in sections 4.4-4.6. These lead to a comparison and evaluation of alternative elimination algorithms in section 4.7. In section 4.8, the selected redundancy elimination algorithm is formally applied within the framework. It is demonstrated through controllability assessments of a nonlinear dynamic chemical process with several fixed alternative control structures. Finally, section 4.9 summarises and concludes the chapter.

4.2 Redundancy Problem within the Dynamic Operability Framework

Chapter 3 presented the inclusion of the Output Controllability Index (Vinson and Georgakis, 2000) into the original Dynamic Operability Framework (DOF), (Bahri, 1996) for controllability assessment of regulatory and fixed structure processes. In regulatory processes, the disturbances and uncertainties $\boldsymbol{\theta}$ perturb the process from the assigned set points, and the deviations may violate the operational constraints. To quantify the effects and the possible violations, the disturbances and uncertainties, the measured process dynamics, and the feasible operating space, are each represented in a multi-dimensional polyhedron. Those are the Expected Disturbance Space (EDS), the

Achievable Operating Space due to disturbances and uncertainties (AOS_θ), and the Desired Operating Space (DOS), respectively. The output controllability index quantifies the extent of process dynamics $\mathbf{w}(t)$ within the feasible operating space. If the collection of all the dynamics is contained within AOS_θ, and the convex-hull of AOS_θ is completely inside DOS, then the process dynamics is completely flexible. Furthermore, several completely flexible processes can be ranked further by the multi-dimensional volume of AOS_θ, which represents the process controllability.

The dimension of EDS is equal to the number of disturbances and uncertainties n_θ. The number of disturbance and uncertainty combinations within EDS is, indeed, infinite. Nevertheless, a number of samples may be assumed to represent EDS adequately. In this study, EDS is discretised into uniform grids across the variables space, and these points relate to the set of disturbance and uncertainty combinations $\boldsymbol{\theta}$. Then the process dynamics due to the combinations are simulated to develop the AOS_θ.

The dimension of AOS_θ is equal to the number of measured output variables being evaluated against the operational constraints, which is n_w. Accordingly, this is also the dimension of DOS. Since the whole dynamics is involved, time is considered as an optional dimension of AOS_θ and DOS. The inclusion quantifies the process dynamics as explained earlier in this section. The exclusion leads to a flexibility mode, where AOS_θ covers the extreme points in output space.

Either at flexibility or controllability mode, the convex-hull vertices of AOS_θ are related to the critical disturbance and uncertainty combination $\boldsymbol{\theta}^k$. This is obtained by projecting the vertices back to EDS at the DOF inner levels (see section 3.6). This step is important because $\boldsymbol{\theta}^k$ represents the minimum set of disturbance and uncertainty combinations to be considered by the subsequent outer levels, which happens to cause the worst possible dynamics.

The computation of AOS_θ convex-hull in this study relies on the availability of the multi-dimensional convex-hull computation software. Such software is the product of an emerging research field of computational geometry since early 1990s. This relatively new occurrence explains the limited number of software, as well as the supported computer platform, computation features, computation speed and documentation. Based on the suitability of the last four conditions, the Qhull (Barber *et al.*, 1996) is selected for convex-hull computation within the framework.. Since Qhull is a C console program, while the Dynamic Operability Framework is developed in MATLAB with Microsoft Windows platform, several MATLAB interface functions have been developed to access the required computation in Qhull.

Qhull does not automatically handle the degeneration (Barber *et al.*, 1996) or redundancy problem, which is the problem where the estimated convex-hull dimension is less than the dimension of data input. Consider there are data for two variables, which are supposed to fill a two-dimensional plane in variable space. If these variables are collinear, their data would form a line segment instead, showing that only one dimension is required to represent both variables. This is a degeneration problem, and either one of the variables is redundant. If the dynamic profiles of collinear variables are monotonic and equally spaced, Qhull immediately detects the degeneracy problem and aborts the convex-hull computation. If the profiles are not monotonic, the Qhull may overlook the collinearity and carry on the convex-hull computation. However, the degenerated dimension then becomes an extremely thin two-dimensional space, containing several small line-segments, which are practically coplanar.

As an illustration of redundancy problem, consider variables A and B in Table 4.1. The variable A contains random points between 0–1, and B = 10A, hence it is collinear with A. As shown in Figure 4.1, Qhull returns six convex-hull vertices, not [0.860,8.600] and [0.151,1.508], which are the boundary points. The returned convex-hull size is 3.6871×10^{-5}, which is the two-dimensional area, not the length of the line segment, which is 7.127.

Table 4.1 Redundancy Problem

	1	2	3	4	5	6	7	8	9	10
A	0.193	0.682	0.302	0.541	0.150	0.697	0.378	0.860	0.853	0.593
B	1.934	6.822	3.027	5.416	1.508	6.979	3.783	8.600	8.536	5.935

= data points
= convex-hull vertices

This is a case of an overlooked redundancy problem due to non-monotonic nature of the input data. When the data is sorted, Qhull immediately detects the problem, returns an error message and aborts the convex-hull computation.

Within the Dynamic Operability Framework, the redundancy problem occurs at the computation of AOS_θ convex-hull when $\mathbf{w}(t)$ includes several collinear dynamic profiles. Since the flexibility index based on the convex-hull computation is a principal specification within the framework, the disruption of AOS_θ convex-hull computation is a major obstacle.

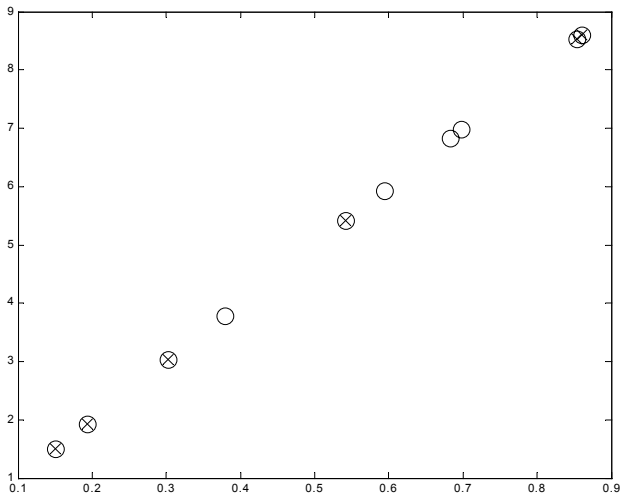


Figure 4.1 Redundancy Problem

O = Data Points, × = Convex-Hull Vertices

The critical disturbance extraction θ^k is also affected by the redundancy problem, because the convex-hull returns coplanar points related to non-critical disturbance combinations. This leads to excess combinations to consider within the framework, as well as the increase in the computational time.

From other point of view, collinearity also indicates measurement redundancy, where the measured value of one variable can be directly determined by measurement of other variables. For a given set of measurement variables, different processes or controller structures may generate different redundant measurements (see section 3.7). Since different process properties have different degrees of measurement difficulties, the redundancy translates to related costs. For instance, pressure sensors are generally cheaper than density sensors. Therefore, if the density of a given process is found redundant, and can be represented by pressure measurement instead, the measurement cost can be reduced. Therefore, the detection and elimination of redundant measurement may lead to significant saving of measurement costs.

These facts outline the importance of the analysis and elimination of the redundant subset of $\mathbf{w}(t)$ within the Dynamic Operability Framework, as well as for other purposes. The elimination procedure is based on the detection of collinearity between variables and the selection of the appropriate variables to construct the AOS_θ. The information sources for the procedure is the dynamic data. This information is readily available, since dynamic simulation is performed intensively within the framework, starting from the first inner level. Moreover, this information can be utilised without any other alterations in the process model.

For redundancy elimination purposes, there are three potential methods to be reviewed in the following sections, namely the Cross-Correlation Analysis, the Principal Component Analysis (PCA) (Jolliffe, 1986) and the Stepwise Collinearity Diagnosis (SCD) (Brauner and Shacham, 2000). The characteristic of each method is analysed, and the suitable features are adapted for redundancy analysis and elimination procedure within the framework.

Further, there are three issues to address during the development of redundancy analysis and elimination procedure. The first issue is to develop the strategy to detect and eliminate the redundancy. If a redundancy is detected, the variables should be properly grouped. Then, one from each group should be selected to construct AOS_θ and DOS, leading to the extraction of critical disturbances. The second issue is to address the effects of the inclusion of time in the redundancy analysis. Time is an independent variable, and therefore can never be redundant. However, its effect on disturbance analysis and elimination is not currently clear, and therefore addressed during the development. The third issue is to address the effects of the redundancy elimination to the quantification of process controllability. These three issues are addressed gradually throughout this chapter.

4.3 Cross Correlation Analysis

The cross-correlation among several variables measures the extent of their collinear relationship. It is a further derivation of covariance, which measures the tendency of each pair of variables to linearly vary together, or co-vary. Consider a set of multi-variable samples in matrix **W**, with size $n_t \times n_w$, where n_t is the number of samples, and n_w is the number of measured variables. The variables are also represented in the column vectors $\mathbf{w}_1$, $\mathbf{w}_2$, …, $\mathbf{w}_{nw}$, respectively. The covariance matrix **R** of these variables is defined as follows:

$$\mathbf{R} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n_w} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n_w} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n_w 1} & \sigma_{n_w 2} & \cdots & \sigma_{n_w n_w} \end{bmatrix} \tag{4.1}$$

$$\sigma_{ij} = E[w_i w_j] - \mu_{w_i}\mu_{w_j} \qquad i, j = 1, 2, \ldots, n_w \tag{4.2}$$

$$E[w_i] = \sum w_i P_i(w_i) \qquad i = 1, 2, \ldots, n_w \tag{4.3}$$

Here, σ is the covariance between two variables, and μ is the mean of the respective variables. The expected value of $\mathbf{w}_i$ is $E[\mathbf{w}_i]$ defined in (4.3), where P_i is the marginal probability distribution of $\mathbf{w}_i$. Note that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are independent, then $\sigma_{12} = 0$.

Therefore, the correlation matrix $\mathbf{R}^*$ is defined as follows:

$$\mathbf{R}^* = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n_w} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n_w} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n_w 1} & \rho_{n_w 2} & \cdots & \rho_{n_w n_w} \end{bmatrix} \tag{4.4}$$

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} \qquad i, j = 1, 2, \ldots, n_w \tag{4.5}$$

The element ρ is the correlation coefficient, where $|\rho_{ij}| \leq 1$. If the absolute value of ρ_{ij} is close to 1, it means that the two variables are highly collinear. The collinearity has the same direction if ρ_{ij} is positive, and an opposite direction if ρ_{ij} is negative. The square of ρ_{ij} is the percentage of collinear variation between variables. For instance, an $\rho_{ij} = 0.5$ represents 25% collinearity.

It is important to highlight that the cross-correlation analysis is suitable for linear relationship only, and the relationship is never assumed causal. The computation of both **R** and **R*** are common procedure in many mathematical or simulation software, such as MATLAB. More importantly, the extent of collinearity captured within the cross-correlation matrix is a prospective basis for redundancy analysis, and this is further explored in section 4.6. Both covariance and correlation matrices are also utilised for further data processing and dimensional reduction, for instance, by the principal component analysis (PCA) described in the next section.

4.4 Principal Component Analysis

The Principal Component Analysis (PCA, Jolliffe, 1986) is widely known and used as one method of multi-variable data analysis. The motivation behind this method is that a set of multi-variable data could contain a substantial set of variable with a large inter-correlation structure. Capturing this correlation may result in a subset containing only the non-correlated **Principal Components** (PCs). These components are ordered, such that the first few retain the most of the variation present in the original data set. The resulting subset could be easier to analyse and interpret. Therefore, PCA serves as pre-processing of many statistical and non-statistical processes. However, while the PCs represent one-or-more of the original variables, it should be noted that PCA is not a one-to-one transformation. Therefore, the inverse transformation is not possible (Jolliffe, 1986).

4.4.1 Computation Procedure

Consider a set of multi-variable samples in matrix **W**. It is a matrix with size $n_t \times n_w$, where n_t is the number of samples, and n_w is the number of measured variables. The idea of PCA is to determine j non-correlated principal components (PCs) $\boldsymbol{\phi}$, which represent the variation in n_w variables. The definition of the PC is given as follows:

$$\boldsymbol{\phi} = \mathbf{S}'\mathbf{W} \tag{4.6}$$

Here, **S** is pre-multiplier matrix with size $n_w \times n_j$. The above relationship can be rewritten using the column vectors of $\boldsymbol{\phi}$, which are $\phi_1, \phi_2, \ldots, \phi_j$, the column vectors of **W**, $w_1, w_2, \ldots, w_{nw}$, and the column vectors of **S**, $s_1, s_2, \ldots, s_j$, and the elements $s_{11}, s_{12}, \ldots, s_{jnw}$, as follows:

$$\begin{aligned}
\phi_1 &= s_{11}w_1 + s_{12}w_2 + \ldots + s_{1nw}w_{nw} = \mathbf{s}_1'\mathbf{W} \\
\phi_2 &= s_{21}w_2 + s_{22}w_2 + \ldots + s_{2nw}w_{nw} = \mathbf{s}_2'\mathbf{W} \\
&\vdots \\
\phi_j &= s_{j1}w_1 + s_{j2}w_2 + \ldots + s_{jnw}w_{nw} = \mathbf{s}_j'\mathbf{W}
\end{aligned} \tag{4.7}$$

The vectors s_i, i = 1, 2, …, j, are chosen such that each ϕ_i contains the maximum variation subject to the following constraints.

$$\mathbf{s}_i'\mathbf{s}_i = 1 \tag{4.8}$$

$$\mathbf{s}_i'\mathbf{s}_k = 0 \quad (i \neq k) \tag{4.9}$$

The aim in general is to find s_i such that that the most variation in **W** is covered by ϕ_1, ϕ_2, …, ϕ_j, where $j<n_t$. To assist the computation, **W** is considered to have a symmetric covariance matrix **R**. Using **R**, the variance of ϕ_i is given as follows:

$$\sup_{\phi} \frac{1}{n_t} \sum_{i=1}^{n_t} \phi_i^2 = \max_{\mathbf{s}_i} \mathbf{s}_i'\mathbf{R}\mathbf{s}_i \tag{4.10}$$

The conditions (4.8) - (4.10) can be solved using the Lagrange multiplier approach in maximising the following function:

$$\mathbf{s}_i'\mathbf{R}\mathbf{s}_i - \lambda_i(\mathbf{s}_i'\mathbf{s}_i - 1) \tag{4.11}$$

It is solved by taking partial differentiation with respect to $\mathbf{s_i}$ as follows:

$$2\mathbf{R}\mathbf{s}_i - 2\lambda_i\mathbf{s}_i = 0 \tag{4.12}$$

It leads to the following eigen-structure of **R**:

$$(\mathbf{R} - \lambda_i\mathbf{I})\mathbf{s}_i = 0 \tag{4.13}$$

$$\left|\mathbf{R} - \lambda_i\mathbf{I}\right| = \begin{vmatrix} r_{11} - \lambda_i & r_{12} & \cdots & r_{1n_w} \\ r_{21} & r_{22} - \lambda_i & \cdots & r_{2n_w} \\ \vdots & \vdots & & \vdots \\ r_{n_w1} & r_{n_w2} & \cdots & r_{n_wn_w} - \lambda_i \end{vmatrix} = 0 \tag{4.14}$$

The characteristic equation of **R** is a polynomial of degree n_w, which is obtained by expanding the determinant of (4.14), and solving for the roots λ_j, j = 1, 2, …, n_w. Specifically, the eigen-vector ϕ_1 associated with the eigen-value λ_1 is the first principal component. After obtaining these values, the computation (4.13)-(4.14) is repeated until all n_w eigen-values are computed. Finally, all of the eigen-values are combined into a diagonal matrix $\mathbf{D}_\phi$.

In summary, the problem of finding **S** in terms of eigen-structure of **R** is as follows:

$$\mathbf{RS} = \mathbf{SD}_\phi \tag{4.15}$$

$$\mathbf{S'S} = \mathbf{SS'} = 1 \tag{4.16}$$

$$\mathbf{S'RS} = \mathbf{D}_\phi \tag{4.17}$$

Here, **S** is the PC score matrix, which serves as the pre-multiplier of the orthogonal transformation in (4.6). The matrices **R** and $\mathbf{D}_\phi$ are the covariance matrix of **W** and the diagonal matrix comprising the eigen-values of **S**, respectively (Jolliffe, 1986).

It is also common to normalise the transformation (4.6) so that the PCs have zero means and unit variances. So far, the principal components (PCs) already have zero means, but not the unit variances. Therefore, PCs with unit variance can be further computed as follows:

$$\tilde{\boldsymbol{\phi}} = \mathbf{D}_\phi^{-1/2}\boldsymbol{\phi} = \mathbf{D}_\phi^{-1/2}\mathbf{S'}\tilde{\mathbf{W}}' = \mathbf{B'}\tilde{\mathbf{W}}' \tag{4.18}$$

$$\mathbf{B} = \mathbf{SD}_\phi^{-1/2} \tag{4.19}$$

The matrix **B** is the 'PC score matrix', $\tilde{\boldsymbol{\phi}}$ is the unit variance PCs, and $\tilde{\mathbf{W}}$ is the unit variance data matrix..

The eigen-values in $\mathbf{D}_\phi$ are sorted to determine the dimension of the PCs. The largest few are typically selected to represent the whole data. Accordingly, the related columns of **B** represent the correlation between PCs and the original variables. Suppose the original data consist of five variables and 2000 measurements. If three largest eigen-values associated with three PCs are retained, then $\tilde{\boldsymbol{\phi}}$ is a 2000×3 matrix. **B** represents the relationship between the normalized original variables with PCs as a 5×3 matrix illustrated in Table 4.2. In this table, b_{ij} is the PC score coefficients of variable i and PC j, where i ranges from 1 to 5 and j from 1 to 3.

Table 4.2 Relation Between Original Variables and Principal Components

Original variables	Principal components		
	1	2	3
1	b_{11}	b_{12}	b_{13}
2	b_{21}	b_{22}	b_{23}
3	b_{31}	b_{32}	b_{33}
4	b_{41}	b_{34}	b_{43}
5	b_{51}	b_{35}	b_{53}

So far, the PCs have been defined based on the covariance matrix **R**. Alternatively, it is also possible to derive the PCs based on either the correlation matrix **R*,** or the weighted covariance matrices, as long as those are symmetric and positive semi-definite. The relative weights of the latter matrix could be chosen to reflect some ideas of the relative importance of the variables. There is a special case where the weights are the square roots of the respective covariance, since it is equal to correlation matrix and leads to the associated correlation PCs. The rest of computational procedures of the correlation PCs are similar to the covariance PCs, using (4.13)-(4.14) (Jolliffe, 1986).

However, the eigen-values and eigen-matrix of correlation matrix have no simple relationship with those of covariance counterparts. Accordingly, the correlation PCs and covariance PCs do not give equivalent information, nor can they be derived directly from each other. In many processes, the correlation based-PCs are preferable, since they can be compared directly with each other, in contrast with the covariance based one. This property is especially useful when dealing with variables with different measurement units. Referring to the CSTR case study in chapter 3, the correlation matrix would enable direct comparison between the temperatures dynamics, ranging from 298 to 315 °K, with the concentration dynamics, ranging from 0 to 0.3. This comparison would be explored further in redundancy analysis in section 4.7.

The covariance matrices, on the other hand, do have the important role when probabilistic inference is required. In this case, the probability analysis is easier if the covariance matrix is used. Furthermore, the determinant of the covariance matrix, which is called 'generalised variance', could be used as a single measure of spread for multivariable data. The square root of the generalised variance for multivariable normal distribution is proportional to the 'volume' of the data set in the reduced dimensional space. This enclosed a fixed proportion of the probability distribution of **W**. For multi-variable normal **W**, the first PCs are the reduced linear functions of **W** whose the contour of joint probability distribution encloses the maximum volume (Jolliffe, 1986).

4.4.2 Singular Value Decomposition in PCA

In practice, the PCA is greatly supported by the singular value decomposition (SVD) function (Jolliffe, 1986). The SVD effectively provides the coefficients and variances of the PCs, as well as the PC score matrix. This function is also readily available in mathematical software, such as MATLAB (Mathworks, 2000).

Given a measurement matrix **W** of dimension $n_t \times n_w$, the SVD computation is defined as follows:

$$\mathbf{W} = \mathbf{ULA}' \tag{4.20}$$

Here, **U** and **A** are $(n_t \times n_w)$, $(n_w \times n_w)$ matrices respectively, each has orthonormal columns so that $\mathbf{U'U} = \mathbf{I}_{nt}$, $\mathbf{A'A} = \mathbf{I}_{nw}$. It follows that n_w is the rank of **W** and the matrix **L** is $n_w \times n_w$ diagonal.

The equation (4.20) shows that SVD provides an efficient method of actually finding PCs. Once **U**, **L**, and **A** which satisfy (4.20) are found, then **A** and **L** provide the eigen-vectors and square roots of the eigen-values of **W'W**, which are the PCs and the covariance respectively. In addition, the SVD also provides the normalised version of PC scores **U**. The columns of **U** are the eigen-vectors corresponding to non-zero eigen-values of **W'W**.

4.4.3 Probabilistic Inference of PCA

The interpretation of PCA so far is descriptive and applicable to all types of variables (continuous, discrete, etc). Nevertheless, there are significant studies related to derivations of approximated probability distributions for the eigen-vectors and eigen-values of a sample covariance matrix (Collins *et al.*, 2000 and the references therein; Jolliffe, 1986; Tipping and Bishop, 1999). This approach is the probabilistic inference about the population of PCs. One main drawback is that all original variables are assumed to have Gaussian distribution, which severely limits the practical values of the application in comparison to the descriptive PCA (Jolliffe, 1986).

The inference PCA assumes that $\mathbf{W}$ is generated from a subset vector ϕ by a linear transformation ($\mathbf{S}$,$\mathbf{W_N}$) plus a noise vector $\mathbf{e}$ as follows:

$$\mathbf{W} = \mathbf{S}\phi + \mathbf{W_N} + \mathbf{e} \tag{4.21}$$

Both noise $\mathbf{e}$ and principal component vector ϕ are assumed as spherical Gaussian:

$$\mathbf{P}(\mathbf{e}) \sim N(0, \sigma_e \mathbf{I}_e) \tag{4.22}$$

$$\mathbf{P}(\phi) \sim N(0, \mathbf{I}_\phi) \tag{4.23}$$

Therefore, $\mathbf{W}$ is also Gaussian:

$$\mathbf{P}(\mathbf{W} \mid \mathbf{S}, \mathbf{W}_N, \sigma_e) \sim N(\mathbf{W}_N, \mathbf{SS}'+\sigma_e \mathbf{I}_e) \tag{4.24}$$

The goal is to estimate the basis vectors $\mathbf{S}$ and the noise variance σ_e from the data set $\mathbf{W}$. The probability of this data is as follows:

$$\mathbf{P}(\mathbf{W} \mid \mathbf{S}, \mathbf{W}_N, \sigma_e) = 2\pi^{-n_t n_w/2} \left|\mathbf{SS}'+\sigma_e \mathbf{I}_e\right|^{-n_t/2} \times \cdots \tag{4.25}$$

$$\exp\left(-\frac{1}{2}\mathbf{tr}((\mathbf{SS}'+\sigma_e \mathbf{I}_e)^{-1}\Omega\right)$$

$$\Omega = \sum_{i=1}^{n_t} (w_i - w_N)(w_i - w_N)' \tag{4.26}$$

Here, w_i and w_N are the column vector of $\mathbf{W}$, representing one observed variable.

The corresponding maximum likelihood estimates for each column are as follows:

$$\hat{w}_N = \frac{1}{n_t}\sum_{i=1}^{n_t} w_i \tag{4.27}$$

$$\hat{\sigma}_e = \frac{\sum_{j=n_\phi+1}^{n_w} \lambda_j}{n_w - n_\phi} \tag{4.28}$$

$$\hat{\mathbf{S}} = \mathbf{L}(\mathbf{R} - \sigma_e^2 \mathbf{I_e})^{1/2} \mathbf{D}_\phi \tag{4.29}$$

Here, $\mathbf{L}$ is an orthogonal matrix containing the top k eigen-vectors of S/n_w, $\mathbf{D}_\phi$ is a diagonal matrix containing the corresponding eigen-values and $\mathbf{R}$ is a covariance matrix (Collins *et al.*, 2000).

The inference studies include the integration of the explicit probability properties into PCA computation. Among others, Tipping and Bishop (1999) demonstrated the explicit account of Gaussian properties to find the principal components using the expectation maximisation (EM) algorithm. This algorithm iteratively maximised the likelihood function, or equivalently minimised the loss function $-\log P(W \mid S, W_N, \sigma_e)$, where $P(W \mid S, W_N, \sigma_e)$ is given in (4.25). Collins *et al.* (2000) extend the PCA interpretation to other exponentially distributed random data (Bernoulli and Poisson), by generalising and minimizing the loss function for use with these probability distributions.

4.4.4 Dimensionality Reduction using PCA in Chemical Processes

Although the descriptive PCA provides wider interpretations, the inferential PCA has been cited more in chemical process engineering. The applications based on the inference PCA span from signal processing, process monitoring, quality controls, process and sensor fault identification and reconstruction, disturbance detection, gross error identification, neural network modeling and product design (Brauner and

Shacham, 2000; Rotem and Lewin, 2000; Singhal and Seborg, 2001; Valle *et al.*, 1999; Wachs and Lewin, 1998; 1999).

Most of those applications involve the determination of dimensionality of the data set. One simple and fast approach is selecting the PCs associated with the assigned threshold of the eigen-values of the PCs (Fukunaga and Olsen, 1971) or cross-correlation matrices (Everson and Roberts, 2000). The variants of these methods include the cumulative percent variance (CPV), the scree test on residual percent variance (RPV), the average eigen value (AE), and the parallel analysis (PA). The CPV and RPV are based on the percent variance captured by the first few or the last few PCs, respectively. The EA accepts all eigen-values higher than the average eigen-value. The scree test looks for the knee point of the RPV plot. However, it could be difficult, since some associated curves might decrease smoothly. The PA computes the PCs of two matrices; one is the original data, and the other is the uncorrelated data matrix with the size equal to the original. The eigen-values of both matrices are then intersected. The values above are considered as information, and those below are the noise. (Valle *et al.*, 1999).

There are also Bayesian methods, which are applied based on the maximum likelihood inference. These methods are iterative and involve weighted integration over probability density functions of each variable. Therefore, they require excessive computation time (Bishop, 1998; Minka, 1999). The variants include the Akaike information criterion (AIC), the minimum description length (MDL), and the imbedded error function (IEF) (Valle *et al.*, 1999).

4.5 Stepwise Collinearity Diagnosis

The stepwise collinearity diagnosis (SCD) procedure (Brauner and Shacham, 2000) has been recently proposed for identifying the collinear and non-collinear subsets. It is

proposed for use either as an alternative or in parallel with PCA. This procedure assumes a perturbed data set similar to (4.21) as follows:

$$\mathbf{W} = \mathbf{W_N} + \mathbf{e} \tag{4.30}$$

The SCD works with normalised variables $\tilde{\mathbf{W}}$. The objective is to provide two lists of variables. The first list is the 'basic set', which represents the maximum subset of non-collinear variables. The second list specifies the variables in each of the identified collinear subsets. The extent of collinearity between variables are measured using the truncation-to-noise (TNR) indicator as follows:

$$\mathbf{TNR_j} = \frac{\left\|\tilde{\mathbf{w}}_\mathbf{j}\right\|}{\left\|\delta\tilde{\mathbf{w}}_\mathbf{j}\right\|} = \left(\frac{\tilde{\mathbf{w}}_\mathbf{j}'}{\delta\tilde{\mathbf{w}}_\mathbf{j}}\frac{\tilde{\mathbf{w}}_\mathbf{j}'}{\delta\tilde{\mathbf{w}}_\mathbf{j}}\right)^{1/2} \tag{4.31}$$

The value $TNR_j > 1$ indicates that the variable $\tilde{\mathbf{w}}_\mathbf{j}$ contains valuable information. In contrast, $TNR_j < 1$ implies that the information is mostly noise.

The SCD procedure consists of consecutive stages. At the beginning of each stage, one of the original variables $\tilde{\mathbf{w}}_\mathbf{j}$ associated with the largest TNR_j is selected to enter the basic set. This variable is referred to as the basic variable $\tilde{\mathbf{w}}_\mathbf{j}^\mathbf{k}$. The remaining are called the non-basic variables $\tilde{\mathbf{w}}_\mathbf{p}$.

The orthogonal of the non-basic variables are then obtained as follows:

$$\tilde{\mathbf{w}}_\mathrm{j}^{\mathrm{k+1}} = \tilde{\mathbf{w}}_\mathrm{j}^{\mathrm{k}} - \tilde{\mathbf{w}}_\mathrm{p}\left(\frac{\left(\tilde{\mathbf{w}}_\mathrm{j}^{\mathrm{k}}\right)\tilde{\mathbf{w}}_\mathrm{p}}{\tilde{\mathbf{w}}_\mathrm{p}'\tilde{\mathbf{w}}_\mathrm{p}}\right) \tag{4.32}$$

The associated error values $\delta\mathbf{x}_j$ and the respective TNR_j of non-basic variables are updated based on the new orthogonals. If the TNR_j is smaller than one, the respective variable is considered collinear with the basic variable, added to the list of collinear subsets and removed from further consideration. The variable with the maximum

TNRj from the remaining group is then selected to become the next basic variable $\tilde{\mathbf{w}}_{\mathbf{j}}^{\mathbf{k}+1}$. Then, procedure continues until all non-basic variables are accounted for.

At this point, a maximum subset of basic variables has been found. Each of the subsets includes some or all of the basic variables and one of the non-basic variables. However, it is still possible that the basic variable associated with a collinear group practically orthogonal to some of the non-basic variables. These variables are removed by performing a regression procedure. This regression provides a linear model of the non-basic variable as a function of some, or all of the basic variables initially included in the subset. The basic variables with low signal-to-noise ratios, and practically orthogonal to the other non-basic variables are removed. The procedure is carried out consecutively until all collinear subsets are free from the orthogonal.

4.6 Analysis of the Existing Methods

In this study, the redundancy analysis and elimination involve the detection of collinearity between the measured dynamic variables and the selection of the functional variables to construct the AOS_θ. The accuracy of the analysis would be indicated by the accuracy of the detection of the critical disturbance combinations related to the AOS_θ convex-hull vertices.

So far, the existing methods provide partial supports for redundancy analysis and elimination. The cross-correlation matrix indicates the collinearity between variables of a given set of measurement data. Nevertheless, it still requires further processing to group the collinear variables together and then to select the functional variables.

The descriptive PCA approach is attractive since it can be applied to a wide range of cases and types of variables. The approach does not assume any probability properties of the data set, therefore does not conflict with the assumed nature of disturbances and uncertainties within the framework at the current stage. The PCA computation can be performed efficiently and quickly using SVD function. However, the resulting PCs still

comprises the combination of all original data, even if only a few dominant PCs are used. Therefore, the PCs themselves do not directly address the variable redundancy problem.

However, the respective PC score matrix (Table 4.2) does indicate the collinearity between variables. Each column of PC score matrix represents the linear combination of the original data that generate the PCs, hence determine the direction of the PCs. The coefficients within this column represent the contribution of each original variable to the respective direction. Collinear variables tend to have similar contribution, hence similar values in the PC score column. Therefore, detecting the similarity among PC score coefficient in each column provides collinearity indication, comparable with those provided by the cross-correlation matrix. Accordingly, further grouping and selection procedure is also required.

The SCD method explicitly addresses the grouping and elimination of redundant variables based on the collinearity. The idea of using the variable with the largest TNR_j as the basic variable, or pivot, for collinearity test is simple and effective. However, the removal of non-collinear data requires several stages of regression upon the subsets until each subset contains only collinear data. If performed in a large set of data, it is less efficient than the computation of cross-correlation matrix and PCA using SVD.

The redundancy analysis and elimination in this study therefore employs the features of the above methods. The efficient computation of cross-correlation matrix and SVD based PCA are widely available. Therefore, the use of both matrices is compared in the analysis. The grouping adopts the pivoting method from SCD. The selection of functional variable from each group, however, has never been addressed before. Therefore, this study selects the variables, based on the distance of their normalised data to the respective normalised constraints, to construct AOS_θ. This leads to the final decision about the elimination quality, which is determined by the extracted critical disturbance.

4.7 Redundancy Analysis

The redundancy analysis is applied on the double CSTR and mixer case study described in chapter 3. There are three fixed control cases being evaluated, namely the open-loop, single-PI control and double-PI controls. All of these cases indicate redundancy problem at the first inner levels. Therefore, the redundancy analysis is based on the associate process dynamics in constructing AOS_θ.

The analysis starts with the projection of the dynamics into two-dimensional planes of the involved variables. The visualisation provides a general idea of the dimension of AOS_θ and leads to justification of the redundancy elimination. Then the variable grouping based on the cross-correlation matrix, covariance and correlation PC scores are compared. The functional variables are selected from each group to construct AOS_θ. Finally, the extracted critical disturbances θ^k related to AOS_θ convex-hull vertices determine the most suitable elimination procedure.

4.7.1 Two-Dimensional Projections of AOS_θ

4.7.1.1 Projection Method

In this analysis, the process dynamics share the same initial conditions, which is the optimum steady state solution. The dynamics due to the member of EDS are simulated and recorded. These dynamics are then normalised with the initial conditions as the mean. Note that the EDS distribution is not symmetric relative to the nominal disturbances. Consequently, the respective normalised dynamics will not be symmetric either.

Figures 4.2-4.4 show the projection of these normalised variables into two-dimensional spaces. The dynamic data in part (a) of the figures consist of 15 combinations of two-dimensional interaction between the dynamic variables, with the time as the only independent variable. The respective steady state data in part (b) consist of only 10 combinations since time is not accounted for.

4.7.1.2 AOS_θ Projection Analysis

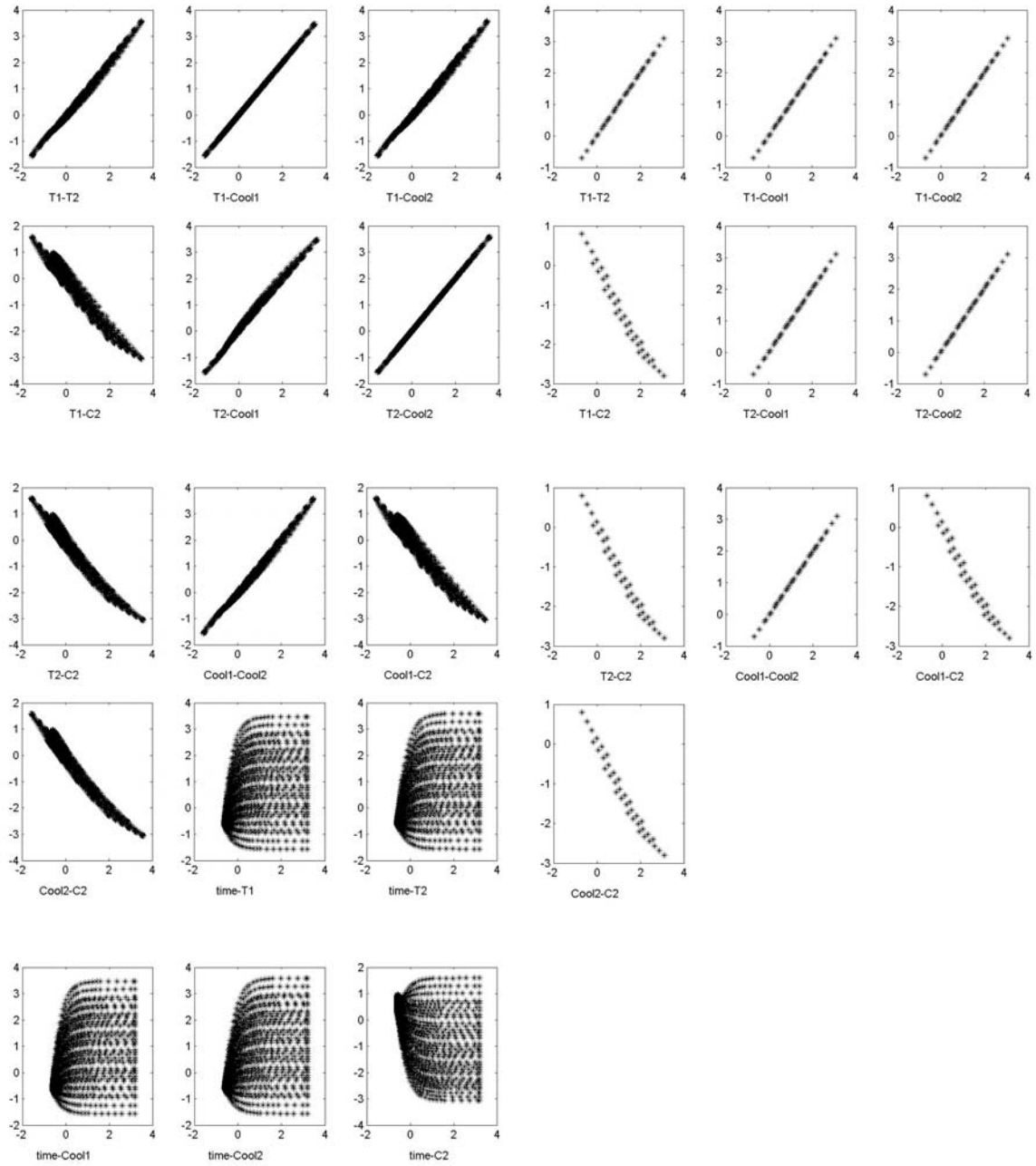
The dynamic data show highly nonlinear relationship between variables, except for the open-loop case. This is the first indication of the unsuitability of performing PCA on these data. In contrast, the steady state data provides clearer interactions.

Both dynamic and steady state projections of the open-loop case in Figure 4.2 show such strong collinearity that T_1, T_2, $Cool_1$ and $Cool_2$ are practically merged into one dimension. Then, C_2 shows slightly weaker collinearity to the other four variables. It is interesting to note that the two-dimensional projections of each four collinear variables and time are almost identical. This indicates variable redundancy, that the dynamics of the variables can be represented by only one of them.

The single-PI control case in Figure 4.3 shows the collinearity between T_2, and $Cool_2$. The steady state projections show further collinearity with $Cool_1$. In this control case, the dynamics of T_1 settles back at the initial condition. Therefore, the variations in the normalised steady state values of T_1 are practically nil (singular). The projections of T_1 with other variables are vertical at T_1 axis, and the shown variations in the figures are within very narrow region, in comparison to the dynamic data.

The dynamic projections for double-PI controls case in Figure 4.4 indicates that all variables are non-collinear, while the steady state projection shows collinearity between T_2 and $Cool_1$, the uncontrolled variables. The steady state residual variations due to the singularities are also shown in the plot of the projection of T_1 and $Cool_2$ to other variables.

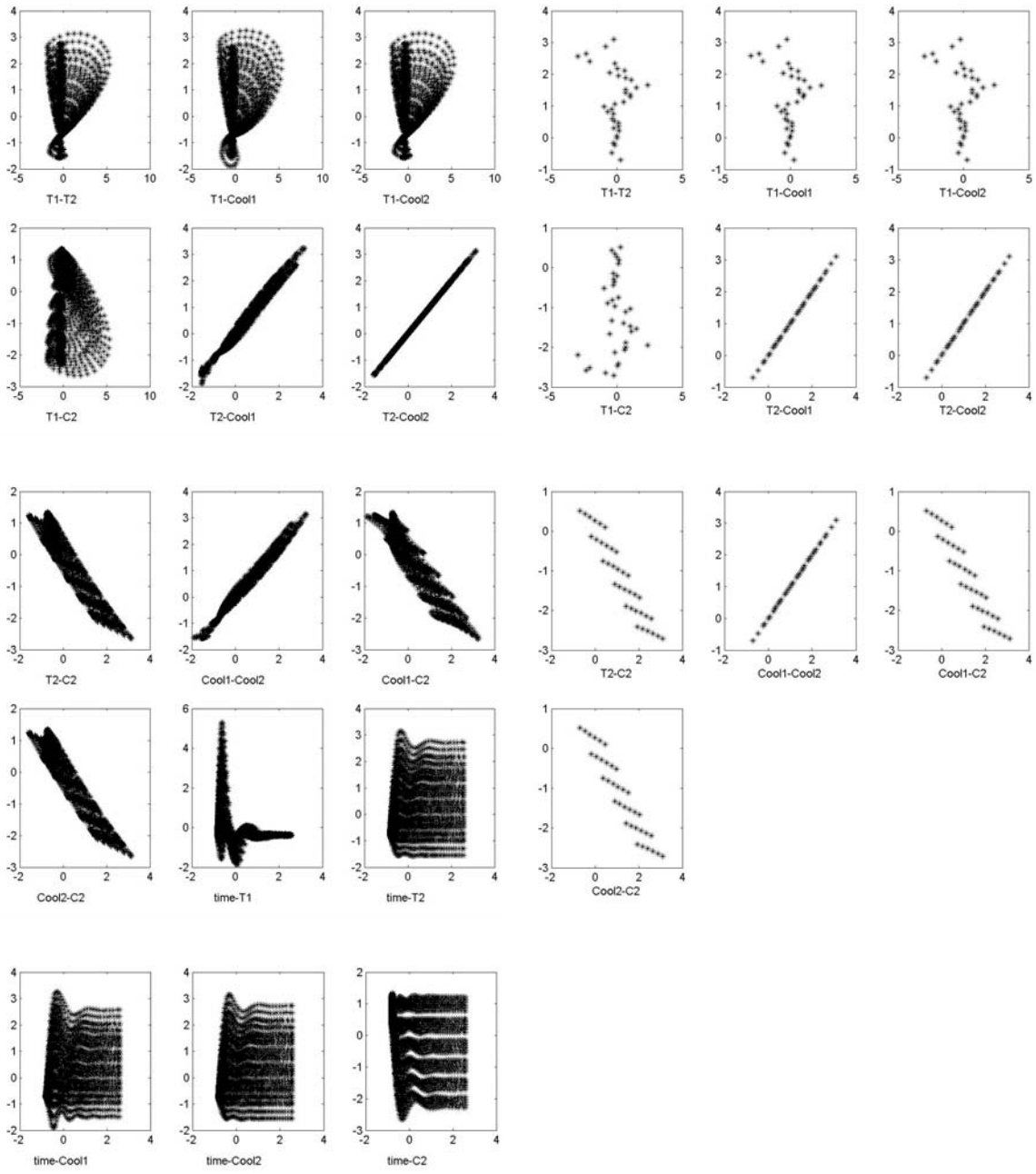
All the figures provide a general idea about the shape and the dimensions of the AOS_θ. All collinear variables form identical two-dimensional projections when combined with other variables. This shows that the reduced dimension AOS_θ is more suitably represented by removal of either one of the collinear or the redundant variable, hence justifies the elimination of redundant variables for AOS_θ construction.



(a) (b)

Figure 4.2 Projection of the Output Values: Open-Loop

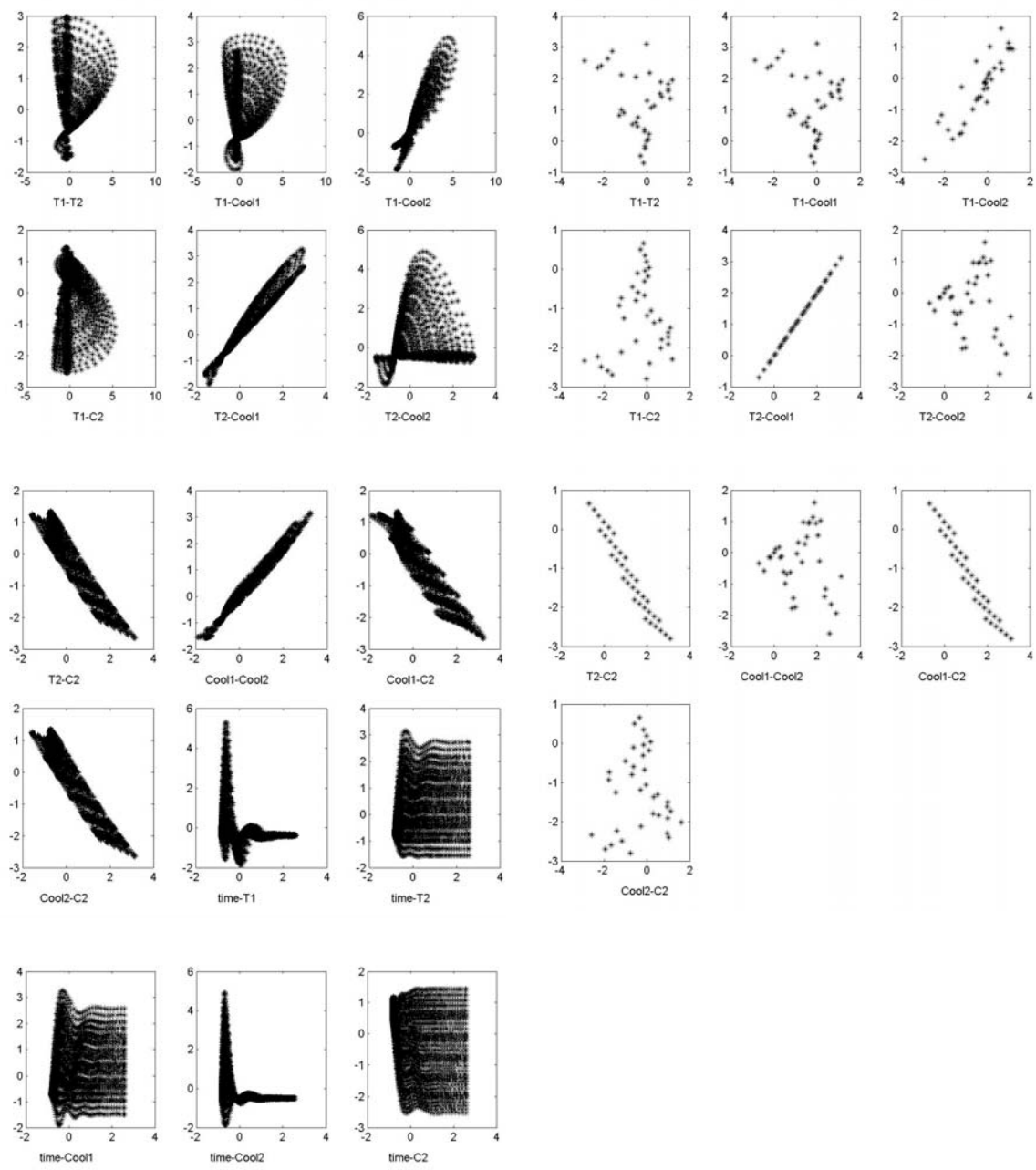
(a) Dynamic Values (b) Steady state Values



(a) (b)

Figure 4.3 Projection of the Output Values: Single PI Control

(a) Dynamic Values (b) Steady state Values



(a) (b)

Figure 4.4 Projection of the Output Values: Double-PI Controls
(a) Dynamic Values (b) Steady state Values

4.7.2 Variable Grouping

In this section, the variables are grouped based on the collinearity between the observed data. Three statistical properties of these variables are compared for this purpose, which are the cross-correlation, the correlation PC and the covariance PC matrices. The 5% tolerance, which is the common threshold in statistical data analysis such as hypothesis testing, are used to decide the collinearity in this study.

The cross-correlation matrix and the correlation PC score matrices of all cases are computed using the normalised data. The covariance PC score, on the other hand, uses the deviation data, derived by taking away the initial conditions from the original data. The PC scores are computed using the SVD function, which is available as MATLAB built-in function. The matrices are given in Tables 4.3 to 4.5. The rows of the matrices are associated with the variables, which are T_1, T_2, $Cool_1$, $Cool_2$, C_2 and time for the dynamic case and the first five for the steady state case. In these tables, the matrix coefficients that are grouped together are indicated by bold typefaces.

4.7.2.1 Cross-Correlation Matrix Grouping

The cross correlation matrix is symmetric. The coefficient values are between zero to one, where the correlation between different variables is indicated by the off diagonal values. Naturally, the diagonal values are equal to one, as they indicate the auto-correlation. These values are used as the basis for variable grouping. Starting from the first column, all variables with the absolute cross-correlation coefficients higher than 0.95, which should include the diagonal variable, are grouped together. The procedure is repeated for the next column. Due to matrix symmetry, if a group containing two or more variables, then there would be the same number of columns representing the group, but with different pivots. Therefore, if the evaluated column contains a pivot that is also a member of a previously detected group, the column is skipped. The evaluation proceeds until all variables or all columns are accounted for.

4.7.2.2 PC Score Matrix Grouping

In PC score matrices, the collinearity is indicated by the absolute similar values between coefficients in the same column. The grouping is initiated from the first column. The highest coefficient is selected as a pivot. The grouping in this column is based on the coefficients that have absolute relative difference less than 5% with the pivot. If all variables have been accounted for, the grouping is stopped. Otherwise, the evaluation carries on to the next column.

4.7.2.3 Grouping Analysis

The cross-correlation and the correlation PC score matrices show similar grouping in most cases in Tables 4.3 to 4.5. On the other hand, the covariance PC score matrices in all cases give the poorest grouping, since the coefficients do not resembles any recognisable pattern, either in the same column, nor in the whole matrix. This problem is due to the different scale of variances between variables. For instance, the achievable dynamic ranges of the temperatures T_1 and T_2 are between 330 to 350K, the amount of cooling $Cool_1$ and $Cool_2$ are between 15 to 25W, and the final concentration C_2 is between 0.23 to 0.32 mol/m^3, depending on the applied control strategy. There is around 1:20 difference between the range of C_2 and T_2, which makes it difficult to compare them based on deviation variables only.

The problem does not occur on correlation PC score and cross-correlation matrices, because both have unit scaled variances that accommodate direct comparison between disturbances. This illustrates the importance of correlation based PCA over the covariance based PCA in chemical processes, as described earlier in section 4.4.1. Therefore, the covariance PCA is not considered further in redundancy analysis.

The correlation PC score and the cross-correlation matrices shown in Tables 4.3 to 4.5 consistently put time in a separate group from other variables at the dynamic analyses. This is logically expected, since time is an independent variable. While the process dynamics in this study are stable, the transients are the complex exponential function of time, and therefore cannot be collinear with time.

Table 4.3 Variable Grouping: Open-Loop

Data	Dynamic data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000 0.998 1.000 0.998 -0.981** 0.542 **0.998 1.000 0.998 1.000 -0.984** 0.566 **1.000 0.998 1.000 0.998 -0.981** 0.542 **0.998 1.000 0.998 1.000 -0.984** 0.566 **-0.981 -0.984 -0.981 -0.984 1.000** -0.576 0.542 0.566 0.542 0.566 -0.576 **1.000**	-0.086 **-0.764** -0.615 -0.003 -0.107 -0.135 -0.071 -0.603 **0.759** 0.012 0.186 -0.146 -0.015 -0.134 -0.108 -0.001 0.609 **0.774** -0.017 -0.147 0.185 0.003 **-0.763** 0.601 0.001 0.005 -0.011 **0.999** 0.000 0.000 **-0.993** 0.114 -0.002 0.000 -0.000 0.000	**0.429** -0.150 -0.256 -0.477 0.456 -0.541 **0.431** -0.115 -0.177 0.519 -0.541 -0.456 **0.429** -0.150 -0.256 -0.477 -0.456 0.541 **0.431** -0.115 -0.177 0.519 0.541 0.456 **-0.427** 0.088 -0.897 0.069 -0.000 0.000 0.277 **0.959** -0.039 -0.031 0.000 0.000
Variable groups	T_1, T_2, $Cool_1$, $Cool_2$, C_2 Time	Time T_1 T_2 C_2 $Cool_2$ $Cool_1$	T_1, T_2, $Cool_1$, $Cool_2$, C_2 Time
Data	Steady state data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000 1.000 1.000 1.000** -0.987 **1.000 1.000 1.000 1.000** -0.987 **1.000 1.000 1.000 1.000** -0.987 **1.000 1.000 1.000 1.000** -0.987 -0.987 -0.987 -0.987 -0.987 **1.000**	**0.754** -0.393 -0.498 0.067 0.159 0.626 0.454 **0.589** -0.218 0.091 0.132 -0.069 -0.087 -0.381 **-0.908** 0.152 0.110 0.143 **0.896** -0.376 -0.005 **-0.789** 0.614 -0.000 -0.000	**0.448** -0.227 -0.498 -0.003 -0.707 **0.448** -0.218 0.502 0.707 -0.003 **0.448** -0.227 -0.498 0.003 0.707 **0.448** -0.218 0.502 -0.707 0.003 -0.445 **-0.896** 0.008 0.000 0.000
Variable groups	T_1, T_2, $Cool_1$, $Cool_2$, C_2	T_1 C_2 T_2 $Cool_2$ $Cool_1$	T_1, T_2, $Cool_1$, $Cool_2$, C_2

Table 4.4 Variable Grouping: Single-PI Control

Data	Dynamic data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000** 0.213 0.190 0.213 -0.162 -0.265 0.213 **1.000** **0.993** **1.000** **-0.955** 0.362 0.190 **0.993** **1.000** **0.993** **-0.952** 0.345 0.213 **1.000** **0.993** **1.000** **-0.955** 0.362 -0.162 **-0.952** **-0.952** **-0.955** **1.000** -0.379 -0.265 0.362 0.345 0.362 -0.379 **1.000**	-0.001 -0.139 **-0.988** 0.065 0.001 0.000 -0.030 -0.564 0.027 **-0.790** -0.006 0.237 -0.041 **-0.801** 0.150 0.578 -0.002 -0.000 -0.007 -0.137 0.007 -0.192 -0.001 **-0.972** 0.000 0.005 -0.001 0.004 **-1.000** 0.000 **-0.999** 0.051 -0.006 0.002 -0.000 0.000	-0.103 **0.768** 0.631 0.030 -0.029 -0.000 **-0.489** 0.043 -0.099 -0.280 0.414 -0.707 **-0.487** 0.037 -0.148 -0.290 -0.810 0.000 **-0.489** 0.043 -0.099 -0.280 0.414 0.707 **0.478** 0.002 0.117 -0.870 0.004 -0.000 -0.216 -0.637 **0.740** -0.021 -0.027 -0.000
Variable groups	T_1 T_2, $Cool_1$, $Cool_2$, C_2 Time	Time $Cool_1$ T_1 T_2 C_2 $Cool_2$	T_2, $Cool_1$, $Cool_2$, C_2 T_1 Time
Data	Steady state data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000** -0.196 -0.198 -0.196 0.133 -0.196 **1.000** **1.000** **1.000** **-0.958** -0.198 **1.000** **1.000** **1.000** **-0.957** -0.196 **1.000** **1.000** **1.000** **-0.958** 0.133 **-0.958** **-0.957** **-0.958** **1.000**	-0.000 **-0.850** -0.512 -0.123 0.000 0.592 -0.353 0.464 0.503 -0.237 **0.793** 0.281 -0.371 -0.393 -0.000 0.144 -0.086 0.113 0.122 **0.972** -0.005 0.258 **-0.609** 0.750 0.000	-0.120 **-0.991** -0.056 -0.001 0.000 **0.499** -0.045 -0.282 0.411 -0.707 **0.499** -0.043 -0.293 -0.814 -0.000 **0.499** -0.045 -0.282 0.411 0.707 **-0.487** 0.109 -0.867 0.008 0.000
Variable groups	T_1 T_2, $Cool_1$, $Cool_2$, C_2	$Cool_1$ T_1 C_2 $Cool_2$	T_2, $Cool_1$, $Cool_2$, C_2 T_1

Table 4.5 Variable Grouping: Double-PI Controls

Data	Dynamic data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000** 0.108 0.191 0.936 -0.076 -0.266 0.108 **1.000** **0.983** 0.088 **-0.975** 0.413 0.191 **0.983** **1.000** 0.156 **-0.959** 0.341 0.936 0.088 0.156 **1.000** -0.047 -0.324 -0.076 **-0.975** **-0.959** -0.047 **1.000** -0.428 -0.266 0.413 0.341 -0.324 -0.428 **1.000**	-0.001 -0.114 **-0.981** -0.145 -0.058 0.002 -0.043 -0.675 0.185 **-0.713** -0.012 0.008 -0.041 **-0.726** -0.018 0.686 0.011 0.001 -0.000 -0.007 -0.054 -0.024 **0.997** -0.046 0.000 0.006 -0.003 0.004 0.046 **0.999** **-0.998** 0.060 -0.007 0.003 0.000 0.000	0.092 **-0.653** 0.277 0.681 -0.082 0.133 **0.554** 0.016 -0.161 -0.113 0.258 0.767 **0.549** -0.045 -0.201 0.145 0.531 -0.595 0.072 **-0.663** 0.210 -0.707 0.023 -0.106 **-0.549** -0.042 0.159 0.032 0.802 0.166 0.277 0.361 **0.888** -0.039 0.043 -0.042
Variable groups	T_1 T_2, $Cool_1$, C_2 $Cool_2$ Time	Time $Cool_2$ T_2 $Cool_1$ T_1 $Cool_2$ C_2	T_2, $Cool_1$, C_2 T_1, $Cool_2$ Time
Data	Steady state data		
Method	Cross-correlation matrix	Covariance PCA	Correlation PCA
Evaluated matrix	**1.000** -0.264 -0.265 0.889 0.199 -0.264 **1.000** **1.000** -0.099 -0.979 -0.265 **1.000** **1.000** -0.101 -0.978 0.889 -0.099 -0.101 **1.000** 0.032 0.199 -0.979 –0.979 0.032 **1.000**	0.000 **-0.761** -0.647 -0.038 0.014 **-0.735** -0.377 0.461 -0.322 0.004 -0.677 0.412 -0.505 0.341 -0.005 0.000 -0.012 -0.006 -0.008 **-0.999** 0.007 0.329 -0.336 **-0.882** 0.005	-0.261 **-0.646** -0.717 -0.014 0.002 **0.553** -0.159 -0.047 -0.401 0.711 **0.553** -0.158 -0.053 -0.415 -0.703 -0.171 **-0.699** 0.693 -0.044 -0.003 **-0.539** 0.208 0.025 -0.815 0.008
Variable groups	T_1 T_2, $Cool_1$, C_2 $Cool_2$	T_2 T_1 C_2 $Cool_2$	T_2, $Cool_1$, C_2 $Cool_2$, T_1

Both dynamic and steady state analyses in the open-loop case shown in Table 4.3 group T_1, T_2, $Cool_1$, $Cool_2$, and C_2 together. This grouping can be explained by the open-loop equations (3.18) and (3.24) in Chapter 3, which show that $Cool_1$ is definitely a linear function of T_1 as $Cool_2$ to T_2. Furthermore, T_2 is also a linear function of T1, which can be derived by combining the equations (3.21) and (3.25).

The collinearity between C_2 and T_2 is weaker, because C_2 contains the exponential function of T_2 (which also means the exponential function of T_1), and depends on C_f, in addition to T_f. Nevertheless, the effect of C_f and T_2 to C_2 are small in comparison to T_f. This small deviation from collinearity is illustrated in Figures 3.5.a and 3.7.a of Chapter 3, as well as in Figure 4.2 of this chapter.

In these figures, the two-dimensional projection of C_2 with other four variables resembles thin rectangular areas. Suppose the width of the rectangle is 5% of the length. Then, the area is only 5% of the length, while the diagonal is 100.12% of the length. This significant difference shows that when collinearity is within 5% tolerance, the corresponding space is more suitably represented by the diagonal, rather than the area.

In terms of the variable ranges, this diagonal is also the diagonal of the space characterised by the ranges of the involved variables. Therefore, if there are n_w collinear variables $\mathbf{w}$, each with a given data set, then the magnitude of the diagonal d_w can be calculated as follows:

$$d_w = \sqrt{\sum_{i=1}^{n_w} \left(\max(\mathbf{w}_i) - \min(\mathbf{w}_i)\right)^2} \tag{4.33}$$

The diagonal d_w can alternatively defined by one of the observed variable in the group ($\mathbf{w}_j$) as follows:

$$d_w = w_j \sqrt{\sum_{i=1}^{n_w} \frac{\left(\max(\mathbf{w}_i) - \min(\mathbf{w}_i)\right)^2}{\left(\max(\mathbf{w}_j) - \min(\mathbf{w}_j)\right)^2}} \tag{4.34}$$

The similar nature of grouping is also shown by the single-PI control case in Table 4.4. In this case, T_1 is consistently separated from T_2, $Cool_1$, $Cool_2$, and C_2. The control on T_1 breaks the collinearity between T_1 and $Cool_1$, since mc_1 is not constant anymore (see equation (3.30) in Chapter 3), and causing the significant offsets of $Cool_1$. The re-inspection of Figure 4.3 shows that while there is no linear analytic relationship between $Cool_1$ and T_2, $Cool_2$ as well as C_2, their dynamic and steady state profiles are reasonably similar. The cross correlation and the correlation PC analyses confirm that the similarities are within the given 5% tolerance, hence group T_2, $Cool_1$, $Cool_2$, and C_2 together. The profiles of T_1, on the other hand, are significantly different to the others, therefore it belong to a separate group.

In the double-PI control, the singularity of the steady state data of the controlled variables T_1 and $Cool_2$ does not cause any computational problem in both analyses. However, the correlation PCA analysis groups the singular variables together, while the cross-correlation analysis does otherwise. Both analyses group the uncontrolled variables together. The re-inspection of Figure 4.3 does indicate weak collinearity between T_1 and $Cool_2$ in both dynamic and steady state cases, but stronger analysis is required to decide this grouping. One possible method is inspecting the resulting critical disturbances, which is discussed in the next section.

4.7.3 Selection of Functional Variables

Since the grouping has still been considered ambiguous, the methods are further evaluated based on the extracted critical disturbances. The extraction is based on the AOS_θ construction. As discussed in earlier sections, the data collinearity leads to redundancy in AOS_θ dimension. Based on the projection and the collinearity analyses, the functional dimension of AOS_θ would depend on the group of the collinear variables.

The remaining problem is how to construct AOS_θ based on these collinear groups. In this study, one functional variable is selected from each group, based on their distance

from the respective normalised constraints. If there is any constraint violation within the group, the variable causing the largest violation is selected. Otherwise, the variable with deviation closest to the constraint is selected. The selected variables are called the functional variables $\overline{\mathbf{w}}$. The functional variables for all control cases are given in Tables 4.6 to 4.11, and are used to construct AOS_{θ}. For accurate size, the magnitudes of these variables are adjusted based on equation (4.32), where $\mathbf{w}_j$ is replaced by $\overline{\mathbf{w}}$. Therefore, the adjusted magnitude of functional variables $d_{\overline{\mathbf{w}}}$ are given as follows:

$$d_{\overline{\mathbf{w}}} = \overline{\mathbf{w}} \sqrt{\sum_{i=1}^{n_w} \frac{\left(\max(\mathbf{w}_i) - \min(\mathbf{w}_i)\right)^2}{\left(\max(\overline{\mathbf{w}}) - \min(\overline{\mathbf{w}})\right)^2}} \tag{4.35}$$

These lead to the extraction of critical disturbance combinations presented in the next section.

4.7.3.1 Extraction of Critical Disturbance Combinations

The functional variables are utilised to construct the AOS_{θ}, and the respective convex-hulls are subsequently computed. The vertices of this convex-hull correspond to the extracted disturbance combinations. Since these may still contain non-critical combinations, the critical combination $\boldsymbol{\theta}^k$ is defined further as the vertices of the extracted disturbance combinations. Therefore, the indication of correct grouping and disturbance extraction is that the disturbances are only detected at the vertices of disturbance space.

Using the above convention, the open-loop and single-PI control cases return the correct $\boldsymbol{\theta}^k$, as shown in Tables 4.6 to 4.9. On the other hand, the double-PI controls only returns the correct $\boldsymbol{\theta}^k$ at the steady state PC score analysis.

Table 4.6 Disturbance Detection: Open-Loop, Dynamic Data

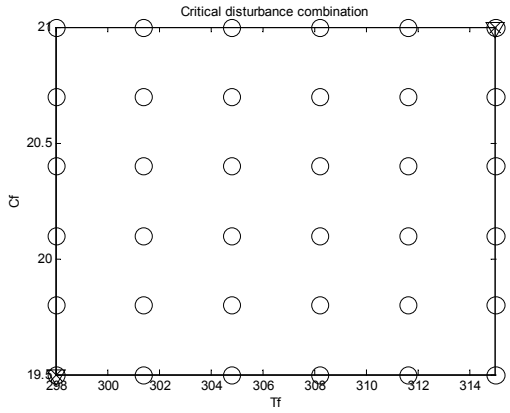
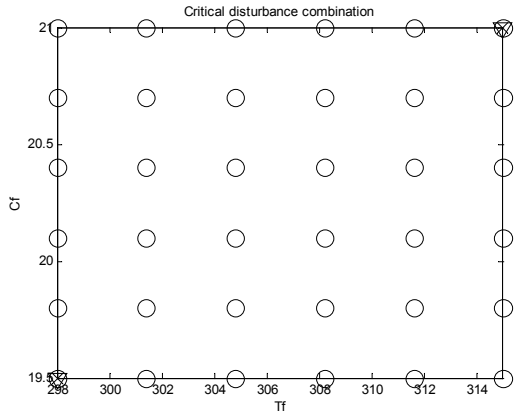
Method	Cross correlation matrix	Correlation PCA
Variable groups	T_1, T_2, $Cool_1$, $Cool_2$, C_2 Time	T_1, T_2, $Cool_1$, $Cool_2$, C_2 Time
Functional variables $\overline{w}$	$Cool_2$ Time	$Cool_2$ time
Disturbance plot o = EDS × = extracted $\nabla = \theta^k$		
θ^k $[T_f, C_f]$	[298, 19.5], [315, 21]	[298, 19.5], [315, 21]

Table 4.7 Disturbance Detection: Open-Loop, Steady state Data

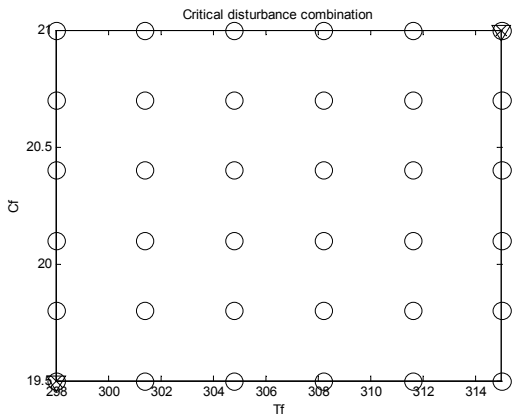
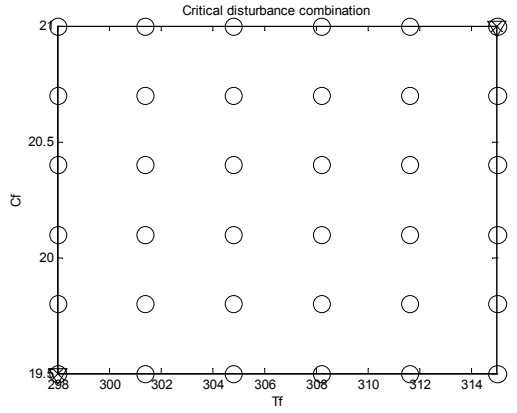
Method	Cross correlation matrix	Correlation PCA
Variable groups	T1, T2, Cool1, Cool2, C2	T1, T2, Cool1, Cool2, C2
Functional variables $\overline{w}$	T1	T1
Disturbance plot o = EDS × = extracted ∇ = θk		
θk [Tf, Cf]	[298, 19.5], [315, 21]	[298, 19.5], [315, 21]

Table 4.8 Disturbance Detection: Single-PI Control, Dynamic Data

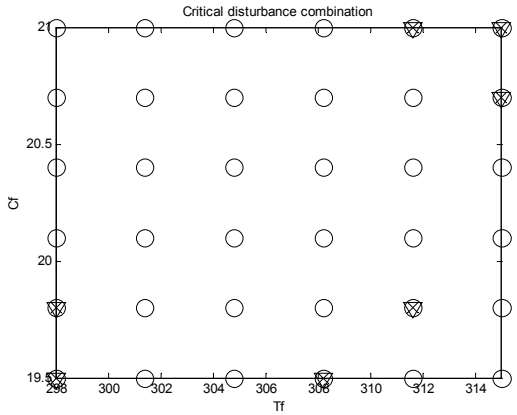
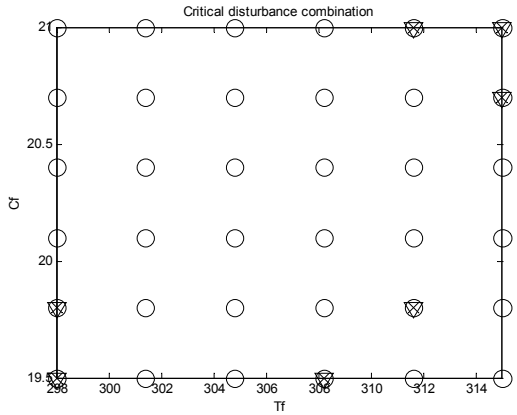
Method	Cross correlation matrix	Correlation PCA
Variable groups	T_1 T_2, $Cool_1$, $Cool_2$, C_2 Time	T_2, $Cool_1$, $Cool_2$, C_2 T_1 Time
Functional variables $\overline{w}$	T_1 $Cool_2$ Time	$Cool_2$ T_1 Time
Disturbance plot o = EDS × = extracted $\nabla = \theta^k$		
θ^k $[T_f, C_f]$	[298, 19.5], [298, 19.8] [308.2, 19.5], [311.6, 19.8] [311.6, 21], [315, 20.7] [315, 21]	[298, 19.5], [298, 19.8] [308.2, 19.5], [311.6, 19.8] [311.6, 21], [315, 20.7] [315, 21]

Table 4.9 Disturbance Detection: Single-PI Control, Steady state Data

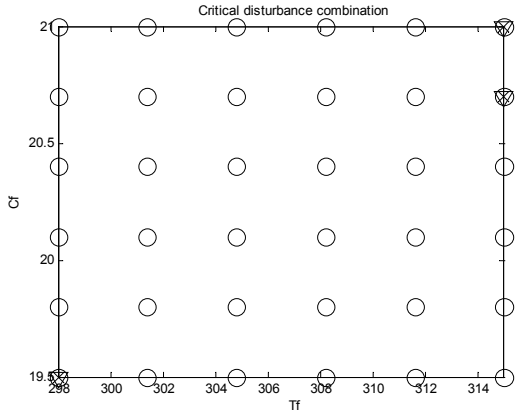
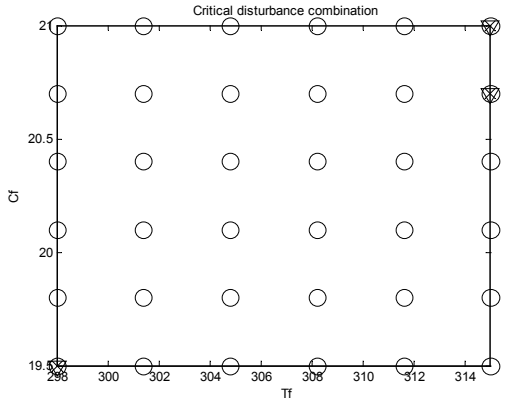
Method	Cross correlation matrix	Correlation PCA
Variable groups	T_1 T_2, $Cool_1$, $Cool_2$, C_2	T_2, $Cool_1$, $Cool_2$, C_2 T_1
Functional variables $\overline{w}$	$Cool_2$ T_1	$Cool_2$ T_1
Disturbance plot o = EDS × = detected $\nabla = \theta^k$		
θ^k $[T_f, C_f]$	[298, 19.5], [315, 20.7], [315, 21]	[298, 19.5], [315, 20.7], [315, 21]

Table 4.10 Disturbance Detection: Double-PI Controls, Dynamic Data

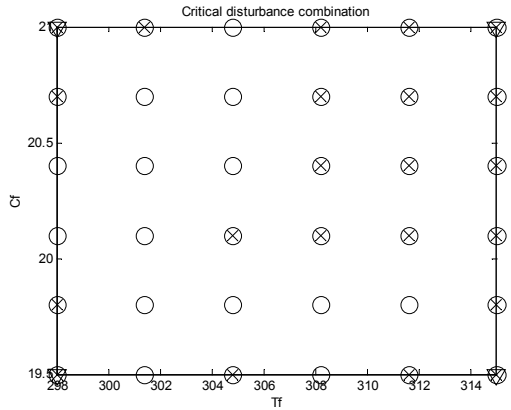
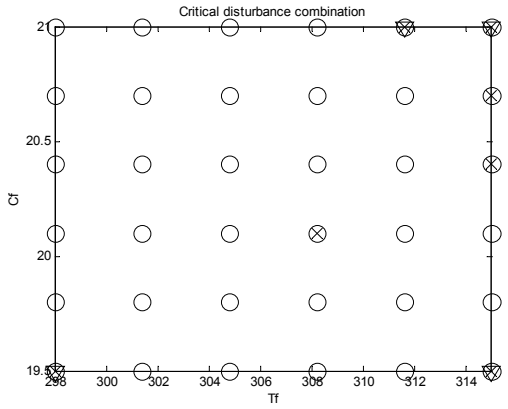
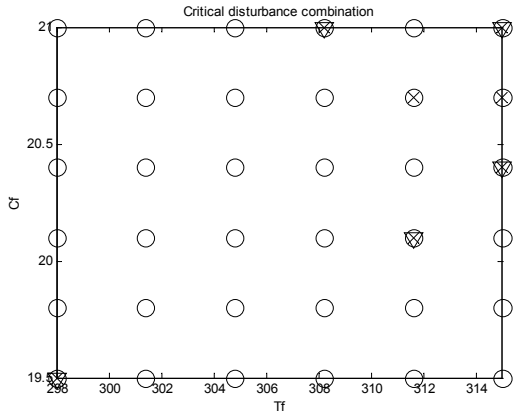
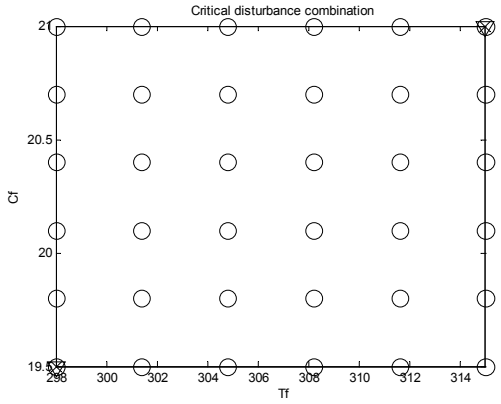
Method	Cross correlation matrix	Correlation PCA
Variable groups	T_1 T_2, $Cool_1$, C_2 $Cool_2$ Time	T_2, $Cool_1$, C_2 T_1, $Cool_2$ Time
Functional variables $\overline{w}$	T_2 T_1 $Cool_2$ Time	T_2 T_1 Time
Disturbance plot o = EDS × = detected $\nabla = \theta^k$		
θ^k $[T_f, C_f]$	[298, 19.5], [298, 21], [315, 19.5], [315, 21]	[298, 19.5], [298, 21], [315, 19.5], [315, 21]

Table 4.11 Disturbance Detection: Double-PI Controls, Steady state Data

Method	Cross correlation matrix	Correlation PCA
Variable groups	T_1 T_2, $Cool_1$, C_2 $Cool_2$	T_2, $Cool_1$, C_2 $Cool_2$, T_1
Functional variables $\overline{w}$	T_1 $Cool_1$ $Cool_2$	$Cool_1$ $Cool_2$
Disturbance plot o = EDS × = detected $\nabla = \theta^k$		
θ^k $[T_f, C_f]$	[298, 19.5], [308.2, 21], [311.6, 20.1], [315, 20.4], [315, 21]	[298, 19.5], [315, 21]

The dynamic cross-correlation and correlation PCA analyses return non-critical disturbances. This problem can be explained by re-inspecting Figure 4.4. This figure shows that the transients at the first few seconds are dense with all critical and non-critical profiles. The density level is such that the associate AOS_θ convex-hull still contains non-critical points, which leads to the extraction of non-critical disturbances. This gives another reason not to use the dynamic data for redundancy analysis, in addition to the difficulty in grouping discussed in section 4.7.2.3.

Grouping the controlled variables in steady state analysis apparently is a better strategy. The cross-correlation grouping leads to non-critical disturbances, while the correlation PCA return the correct ones. Furthermore, the steady state correlation PCA also consistently returns the smallest number of correct critical disturbances at all control cases.

All cases also show that the redundancy analysis is only affected by the dynamic variables, not by time as an independent variable. Therefore, time is not considered in the redundancy analysis, but still used to calculate the controllability indices within the framework. The effects of redundancy elimination to the indices are evaluated further in section 4.8

4.7.4 Realisation of Redundancy Analysis and Elimination Procedure

Based on the above results, the steady state correlation PCA is selected for redundancy analysis and elimination procedure within the Dynamic Operability Framework. The procedure is realised as *pcatest1.m*, an additional function to DOF Toolbox.

At this stage, the function is operated at the inner level. At this level, the process dynamics due to all members of EDS are simulated and recorded. The evaluated steady state values $\mathbf{w}(t_{ss})$ are extracted from these recorded data. This data is sent to *pcatest1.m* along with the initial conditions $\mathbf{w}(0)$, the respective constraints $\mathbf{w}^{\mathbf{L}}$ and $\mathbf{w}^{\mathbf{U}}$, and the collinearity threshold $\boldsymbol{\varphi}$.

The function starts with the normalisation of $\mathbf{w}(t_{ss})$, $\mathbf{w}^{\mathbf{L}}$ and $\mathbf{w}^{\mathbf{U}}$ around $\mathbf{w}(0)$. Then the correlation PC score is computed using the singular value decomposition of the normalised steady state data $\tilde{\mathbf{w}}(t_{SS})$. The resulted PC score is then used for variable grouping using the collinearity tolerance $\boldsymbol{\varphi}$. The functional variables from each group $\overline{\mathbf{w}}^{\mathbf{k}}$ are determined based on the distance from the normalised constraints. The scaling factors **ff**, for calculating the associated $d_{\overline{\mathbf{w}}}$, are given as follows:

$$\mathbf{ff} = \sqrt{\sum_{i=1}^{n_w} \frac{\left(\max(\mathbf{w}_i) - \min(\mathbf{w}_i)\right)^2}{\left(\max(\overline{\mathbf{w}}) - \min(\overline{\mathbf{w}})\right)^2}} \tag{4.36}$$

Finally, the function returns the groups [**w**], the functional variables $\overline{\mathbf{w}}^{\mathbf{k}}$, and the scaling factor **ff**.

The functional variables $\overline{\mathbf{w}}^{\mathbf{k}}$ are used for further extraction of the critical disturbances $\boldsymbol{\theta}^{k}$, which are considered in the subsequent outer level. Within the outer level, the functional variables are used along with time t for computation of the controllability index r-OCI.

4.8 Application of Redundancy Analysis and Elimination Procedure

Finally, the redundancy analysis and elimination using the steady state correlation PCA method function is formally applied within the Dynamic Operability Framework on all fixed control cases. The redundancy elimination is performed in the first inner-level, where the process is subjected to all members of EDS. The screened variables construct the AOS_{θ}. The critical disturbance combinations $\boldsymbol{\theta}^{\mathbf{k}}$ related to AOS_{θ} convex-hull vertices are extracted, then sent to the next outer level.

The outer-level uses $\boldsymbol{\theta}^{\mathbf{k}}$ to find the optimum dynamic operational conditions and the functional variables to compute the controllability index r-OCI. To compute the index,

DOS is constructed using the same functional variables as for AOS_{θ}. The optimum conditions are validated at the second inner-level. At this level, no new critical disturbances are found, and the controllability index is equal to one. Therefore, the framework is completed in the second iteration. The results are shown in Table 4.12, and the critical disturbances are shown in Figure 4.5.

The detected critical disturbances are consistent between the first and the second iterations for all cases. In the single-PI control case, the number of critical disturbances found during the iterations decreases. This is due to process nonlinearity, where the change of operating conditions can change the profiles, as well as the critical disturbances. All critical disturbances are found on the extreme vertices of EDS, which is expected from a monotonic process. The framework computational cost using the redundancy analysis and elimination procedure is improved by 40% in comparison to the manual elimination presented in Chapter 3.

The computation of the GAE and GIAE are significantly affected by the redundancy elimination. Due to process non-linearity, the involved functional variables apparently vary between control cases and between iterations, so do the unit scales. At this point, the scaling factor **ff** provides the required adjustment for the functional variables, such that GAE and GIAE contain the respective original size as close as possible. The theoretical justification to this adjustment is recommended for future work.

The controllability index r-OCI is the volume ratio between two polyhedrons with the same dimensions, therefore it is unaffected by the dimensional reduction. Furthermore, since the functional variables are those closest to, or those violate the constraints, the associate feasibility constraint on r-OCI is still applicable.

The profit variation PVAR captures the variation of the objective function over time within this case study. It is computed using a fixed combination of variables, therefore is not affected by the redundancy elimination.

Table 4.12 The Framework Results with Redundancy Elimination

Case study	Φ ($/hr)	r-OCI	GAE	GIAE	PVAR ($)	w^k	θ^k	CPU time (min)
Dynamic open-loop (iter#1)	90.35	0.184	16.69	1512.75	302.19	T_1	[315,21], [298,19.5]	2.84
Dynamic open-loop (iter#2)	46.86	1.00	14.76	1305.48	189.343	T_1	[315,21], [298,19.5]	
Dynamic closed-loop single-PI control (iter#1)	90.35	0.190	6.58	341.88	168.90	$Cool_2$, T_1	[315,21], [315,20.7], [298,19.5]	3.67
Dynamic closed-loop single-PI control (iter#2)	79.27	1.00	4.20	234.52	147.86	$Cool_2$, T_1	[315,21], [298,19.5]	
Dynamic closed-loop double-PI controls (iter#1)	90.35	0.356	2.10	107.34	21.84	$Cool_1$, $Cool_2$	[315,21], [298,19.5]	3.86
Dynamic closed-loop double-PI controls (iter#2)	87.77	1.00	0.26	14.79	15.10	C_2, T_1	[315,21], [298,19.5]	

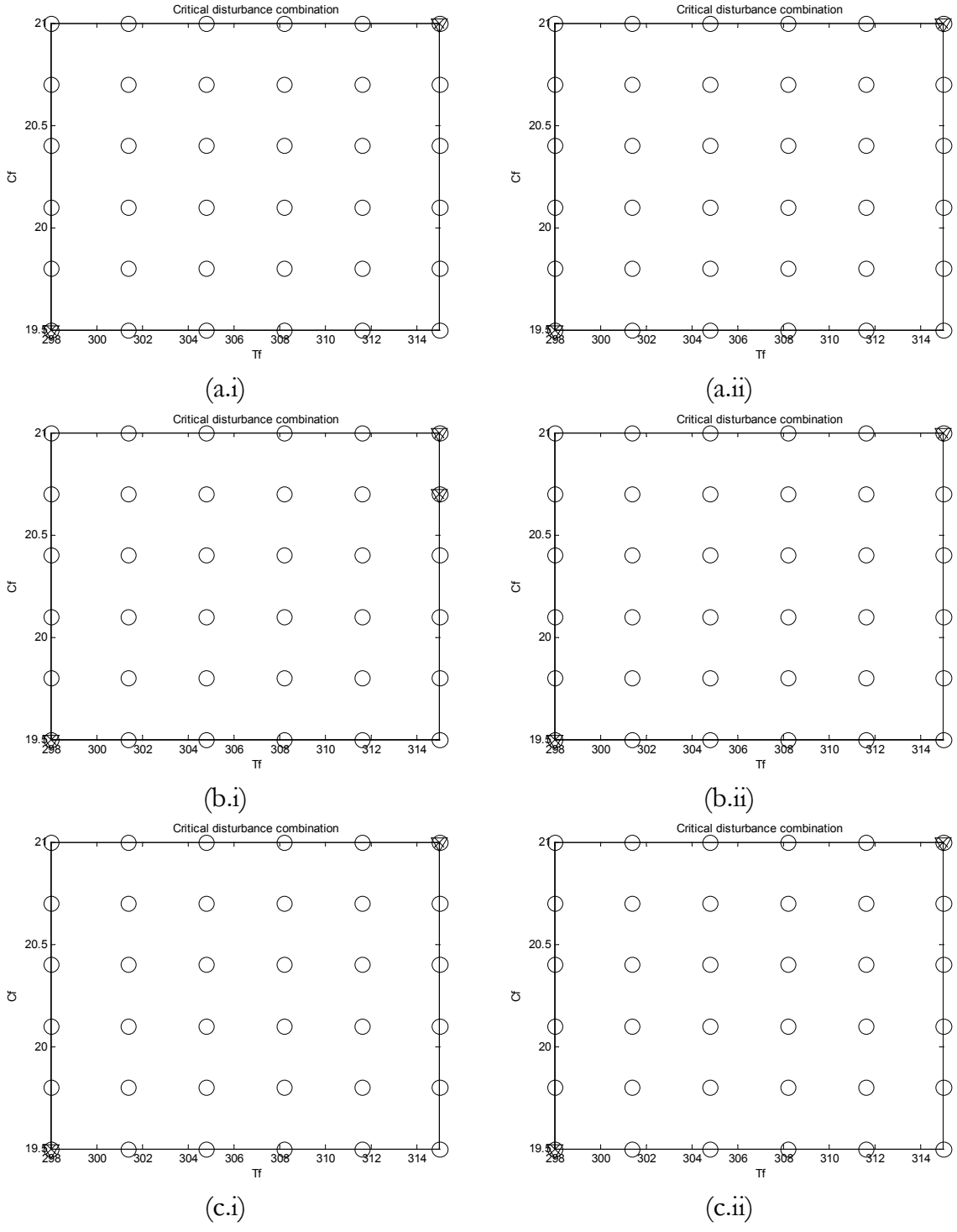


Figure 4.5 Critical Disturbance Combination Found in Both Inner-Levels

(a) Open-Loop (b) Single-PI control (c) Double-PI controls
(i) Iteration 1 (ii) Iteration 2
∘ = EDS Member × = Detected Combination ∇ = Critical Combination

Based on the above conditions, r-OCI and PVAR can be suitably applied in conjunction with the redundancy elimination across several alternative process and controller structures. This feature is further explored in the extension of the Dynamic Operability Framework to cover the automatic selection of feasible process and controller structures.

4.9 Conclusion

A redundancy analysis and elimination procedure for the constrained measurement variables within the Dynamic Operability Framework has been developed and applied for regulatory cases of fixed process and controller structures. The elimination procedure is based on the correlation PCA analysis of process steady state data. It leads to the extraction of the smallest number of critical disturbances affecting the process. Accordingly, the computational times are significantly reduced.

The elimination procedure has been implemented into a MATLAB function, and added to DOF toolbox. The capability of finding different functional measured variables, or equivalently eliminating redundant variables, has been shown for fixed structure cases. It is further explored in the simultaneous assessment of process and controller structures, which is presented in the next chapter.

4.10 Nomenclature

Acronyms		
AOS_θ	:	Achievable Output Space due to disturbances
CSTR	:	Continuous Stirred Tank Reactors
DOF	:	Dynamic Operability Framework
DOS	:	Desired Output Space
EDS	:	Expected Disturbance Space
GAE	:	Generalised Absolute Error
GIAE	:	Generalised Integral Absolute Error
r-OCI	:	Regulatory Output Controllability Index
PC	:	Principal Components
PCA	:	Principal Component Analysis
PI	:	Proportional – Integral controller
PVAR	:	Profit Variation
SCD	:	Stepwise Collinearity Diagnosis
SVD	:	Singular Value Decomposition
TNR	:	Truncation to Noise ratio

Variables and matrices		
A	:	Right unitary orthonormal matrix in SVD computation
B	:	PC score matrix
C_2	:	Product concentration (mol/m^3) of $CSTR_2$
$Cool_1$	:	Amount of cooling (KJ) in $CSTR_1$
$Cool_2$	:	Amount of cooling (KJ) in $CSTR_2$
D_ϕ	:	Diagonal matrix containing the eigen values of **R**
E(•)	:	Expected value of variable •
e	:	Noise matrix
ff	:	Scaling factor of functional variables
I	:	Identity matrix
L	:	Square roots of the eigen values of **W′W**
N	:	Gaussian probability distribution function
n	:	Number of variables or components
P**(•)**	:	Probability of variable •
R	:	Covariance matrix
R*	:	Correlation matrix
S	:	Generalised PC score matrix
$s_1, \ldots, s_n$	:	Column vector of PC score matrix
T_1	:	Temperature (°K) of $CSTR_1$
T_2	:	Temperature (°K) of $CSTR_2$
t	:	Time (sec)
U	:	Left unitary orthonormal matrix in SVD computation

Variables and matrices		
$\mathbf{u}$	:	Vector of manipulated variables
$\mathbf{W}$	:	Set of output/measured variable data
$\mathbf{W_N}$	:	Set of output/measured variable data, assumed with gaussian distribution
$\tilde{\mathbf{W}}$	:	Set of normalised output/measured variable data
$w_1, \ldots, w_n$	:	Column vector of measured data matrix
$\mathbf{w}$	:	Output/measured variables
$\tilde{w}$	:	Column vector of normalised output/measured variable data
$\overline{\mathbf{w}}$	:	Functional variables

Greek letters		
δ	:	Error values in SCD computation
$\mathbf{\Phi}$	:	Objective function
$\boldsymbol{\phi}$	:	Principal components
$\tilde{\phi}$	:	Unit variance principal components
$\phi_1, \ldots, \phi_n$	:	Column vector of principal component matrix
φ	:	Collinearity treshold
λ	:	Lagrange multiplier, the eigen values of $\mathbf{R}$
μ	:	Mean values
θ	:	Disturbances
ρ	:	Cross-correlation coefficients
$\mathbf{\Omega}$	:	Sum of deviation data
σ	:	Variances

Superscripts		
k	:	Critical values
l	:	Lower bounds
u	:	Upper bounds
N	:	Nominal values

Subscripts		
1	:	First CSTR
2	:	Second CSTR
e	:	Noise
f	:	Feed
i, j	:	Arbitrary indices
p	:	Basic variables
ss	:	Steady state, initial condition

4.11 References

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Barber, C. B., Dobkin, D. P. and Huhdanpaa, H. T. (1996). "The Quickhull Algorithm for Convex-Hulls", ACM Transaction on Mathematical Software, 22(4), 469 - 483.

Bishop, C. M. (1998). "Bayesian PCA", Neural Information Processing Systems 11, 382 - 388.

Brauner, N. and Shacham, M. (2000). "Considering Precision of Data in Reduction of Dimensionality and PCA", Computers and Chemical Engineering, 24, 2603 - 2611.

Collins, M., Dasgupta, S. and Schapire, R. E. (2000). "A Generalisation of Principal Component Analysis to the Exponential Family".

Everson, R. and Roberts, S. (2000). "Inferring the Eigen Values of Covariance Matrices from Limited, Noisy Data", IEEE Transactions of Signal Processing, 48(7), 2083 - 2091.

Fukunaga, K. and Olsen, D. (1971). "An Algorithm for Finding Intrinsic Dimensionality of Data", IEEE Transactions on Computers, 20(2), 176 - 183.

Jolliffe, I. T. (1986). "Principal Component Analysis", New York, Springer-Verlag.

Mathworks (2000). "Using MATLAB Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Minka, T. P. (1999). "Automatic Choice of Dimensionality for PCA", 514, MIT Media Lab Vision and Modelling Group, Cambridge, MA 02139.

Rotem, Y. and Lewin, D. (2000). "Assessing the Impact of Parametric Uncertainty of the Performance of Model-Based PCA", ADCHEM 2000.

Singhal, A. and Seborg, D. E. (2001). "Matching Patterns from Historical Data Using PCA and Distance Similarity Factors", American Control Conference, Arlington, VA, USA.

Tipping, M. E. and Bishop, C. M. (1999). "Probabilistic Principal Component Analysis", Journal of the Royal Statistical Society, Series B, 61(3), 611-622.

Valle, S., Li, W. and Qin, J. (1999). "Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods", Industrial and Engineering Chemistry Research, 38(11), 4389 - 4401.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10(2), 185 - 194.

Wachs, A. and Lewin, D. R. (1998). "Process Monitoring Using Model-Based PCA", 5th IFAC Symposium of Dynamics and Control of Process Systems, Corfu, Greece.

Wachs, A. and Lewin, D. R. (1999). "Improved PCA Methods for Process Disturbance and Failure Identification", AIChE Journal, 45(8), 1688 - 1700.

Chapter 5

Process and Controller Synthesis

5.1 Introduction

This chapter presents the extension of the proposed Dynamic Operability Framework to cover the automatic process and controller synthesis. The problem is formulated as a dynamic, semi-infinite mixed-integer nonlinear programming (MINLP) problem. The algorithmic solution of this problem is developed based on the 'Branch and Bound' method, and tailored to accommodate the redundancy problem associated with each process and controller structure candidates. The algorithm is demonstrated on four cases of process and controller synthesis of chemical process superstructures.

The early results of the synthesis problem in this chapter has been published in a conference paper titled "Systematic Controllability Assessment in Dynamic Operability Framework', presented at the Third Inter University Postgraduate Electrical Engineering Symposium, Australia, 2002. The integrated process synthesis and redundancy elimination has also been summarised in a conference paper. The title is "Variable Redundancy Elimination and Automatic Structure Selection within Dynamic Operability Framework", for presentation at the Process System Engineering Symposium in Kunming, P. R. China, January 5th-9th 2004.

The structure of this chapter is as follows: Section 5.2 discusses the problem of process and controller structure selection in chemical processes. It is followed by the formulation of the DOF process synthesis problem in section 5.3, and its development in section 5.4. Section 5.5 demonstrates the application of the proposed framework on four case studies of nonlinear chemical process syntheses. Finally, section 5.6 summarises and concludes the chapter.

5.2 Controllability Consideration in Chemical Process Synthesis

The previous chapters presented the integration of the Output Controllability Index (Vinson and Georgakis, 2000) into the Dynamic Operability Framework (DOF) (Bahri, 1996) for regulatory processes, and the redundancy problem that emerged from the index computation. The associated redundancy analysis and elimination procedure has been developed based on the correlation PCA analysis of the steady state process data. It eliminates the redundant variables, or equivalently finds functional measured variables, which leads to extraction of the critical disturbances affecting the process.

The controllability assessments using this framework have been demonstrated on nonlinear processes with fixed structures. However, while useful comparisons of controllability resulting from different control strategies were presented, they were based on the separate assessments of each process. Furthermore, the possibility of improving the process controllability and profitability by altering process and controller configurations has not been explored. In this chapter, the framework is expanded to accommodate simultaneous assessments of possible configurations within a given process control superstructure. The variable redundancies associated with each structure candidates are also accounted for. The expansion is considered as a case of process synthesis, which is described in the following section.

5.2.1 Chemical Process Synthesis

Due to their inherent nonlinearity, chemical processes present varying dynamic characteristics over different operating conditions and different structures. Consequently, certain dynamic and economic performances may only be achieved by optimising the process and controller structures, in addition to optimising the parameters. At this point, the difficulty arises due to strong interaction between both aspects, as they cannot be isolated and optimised separately. This interaction between structure and parameter in process design, in pursuing the economic and control objectives, is best addressed within process synthesis problems (Nishida *et al.*, 1981).

The synthesis problem is to select the optimum process structure, as well as its design parameters. It involves rigorous analysis of the existence and interconnection of unit operations, as well as the sizes and parameters of the components. The former clearly implies making discrete decisions, while the latter implies making a choice from among a continuous space (Grossmann, 1990). The analysis ultimately determines the quality of process controllability and profitability in presence of disturbances and uncertainties.

Over the last three decades, process synthesis problems have been investigated using the heuristic, the physical insights, and the mathematical programming approaches (Grossmann, 1990). The physical insights yield in many essential design procedures and guidelines, as well as novel representations (Fisher *et al.*, 1988; Wolff *et al.*, 1992; Zhao and Skogestad, 1997). Nevertheless, the development of automated synthesis procedures are mostly based either on the knowledge based heuristic, or on mathematical programming approach (Bahri, 1996).

In a continuation of the work in previous chapters, this chapter focuses on the mathematical programming approach to process synthesis. This approach translates the process synthesis problem into the process control parameters and structures specification problem to satisfy the controllability and profitability objectives. It leads to two distinctive decisions. The first decision is about the operational parameter, for

instance, the size and operating conditions of the involved unit operations, which are continuous in value. The second decision is about process control structure. This may include the existence of a unit, stream or interconnection, as well as the corresponding operations. These may involve 'yes-no', 'if-then' and other logic decisions, as well as selection of several distinct operational conditions. Hence, the decision values are discrete. The consideration of these type of decisions gives rise to a Mixed Integer Nonlinear Programming (MINLP) problem.

5.2.2 MINLP Formulation and Solution Procedure

The common MINLP formulation used for process synthesis is as follows:

$$\min_{\bar{\mathbf{z}},\mathbf{y}} \Phi(\bar{\mathbf{z}},\mathbf{y},\boldsymbol{\theta},\bar{\mathbf{w}}) \tag{5.1}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}},\mathbf{y},\boldsymbol{\theta},\bar{\mathbf{w}}) = \mathbf{0} \qquad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{g_j}(\bar{\mathbf{z}},\mathbf{y},\boldsymbol{\theta},\bar{\mathbf{w}}) \le \mathbf{0} \qquad \mathbf{j} \in \mathbf{I}$$

$$\bar{\mathbf{z}} \in \bar{\mathbf{Z}} = \{\bar{\mathbf{z}} : \bar{\mathbf{z}}^{\mathbf{l}} \le \bar{\mathbf{z}} \le \bar{\mathbf{z}}^{\mathbf{u}}\}$$

$$\mathbf{y} \in \{0,1\}$$

$$\boldsymbol{\theta} \in \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\}$$

$$\bar{\mathbf{w}} \in \bar{\mathbf{W}} = \{\bar{\mathbf{w}} : \bar{\mathbf{w}}^{\mathbf{l}} \le \bar{\mathbf{w}} \le \bar{\mathbf{w}}^{\mathbf{u}}\}$$

Here, $\boldsymbol{\Phi}$ is the process objective, which is the function of independent continuous design variables $\bar{\mathbf{z}}$, discrete design variables $\mathbf{y}$, dependent variables $\bar{\mathbf{w}}$ and disturbances $\boldsymbol{\theta}$. The continuous design variables $\bar{\mathbf{z}}$ represent continuous design values such as flow rates, compositions, temperatures and equipment sizes. The discrete design variables $\mathbf{y}$ may take several different discrete values. Nevertheless, many cases can be reformulated such that only binary values of $\mathbf{y}$ are required. In these cases, the formulation assigns y=1 if a certain condition is applicable, and y=0 otherwise. Other process specifications can be represented in $\mathbf{E}$ equalities in the set $\mathbf{h_i}$, such as process models, and in $\mathbf{I}$ inequalities in the set $\mathbf{g_j}$, such as operational limits. As explained in Chapter 2, the nature of these specifications may involve process dynamics, nonlinearity and uncertainties.

In general, a process synthesis involving the MINLP approach consists of three major steps. Firstly, a superstructure containing all process structure candidates is postulated. Then the superstructure is formulated as a MINLP problem (5.1). Additional considerations such as process dynamics, nonlinearity and uncertainties may be involved, and may require extended formulations. Finally, the optimum process parameters and structure is extracted from the superstructure by solving the MINLP problem along with its extended features (Bahri, 1996; Mohideen *et al.*, 1996; Schweiger and Floudas, 1997).

Superstructure development is a common task in network of homogeneous processes, such as heat exchanger network (HEN), reactor network, and distillation sequences. Once the interconnection patterns amongst the unit operations are identified, the MINLP formulation of the whole network can be developed compactly using the format given in (5.1). Unfortunately, this may be difficult in heterogeneous system, such as a plant wide flowsheet. Since certain discrete patterns may not exist, each unit raises its own MINLP problem and contributes to large overall MINLP formulations (Bahri, 1996).

The discrete decisions in MINLP cases can be translated into constraints involving binary (0-1) variables. Typical translation methods include the Big-M approach and the disjunctive approach (Raman and Grossmann, 1994). The disjunctive approach expresses a logic condition using the Conjunctive Normal Form (CNF) as follows:

$$\vee\begin{bmatrix} \mathbf{Y}_{ik} \\ \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}, \bar{\mathbf{w}}) \le 0 \end{bmatrix} \qquad \boldsymbol{\Omega}(\mathbf{Y}_{ik}) = \text{true} \tag{5.2}$$

In CNF, $\mathbf{Y}_{ik}$ are the logic variables that enforce $\mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}, \bar{\mathbf{w}}) \le 0$ when the logic relationship $\boldsymbol{\Omega}(\mathbf{Y}_{ik}) = \text{true}$ is satisfied. For example, to use $\mathbf{L}_1 \le \mathbf{z} \le \mathbf{U}_1$ if $\mathbf{y}$=1 and $\mathbf{L}_2 \le \mathbf{z} \le \mathbf{U}_2$ if $\mathbf{y}$=0, the expression is as follows:

$$\vee\begin{bmatrix} \mathbf{y} \\ \mathbf{L_1} \le \mathbf{z} \le \mathbf{U_1} \end{bmatrix} \vee \begin{bmatrix} \neg\mathbf{y} \\ \mathbf{L_2} \le \mathbf{z} \le \mathbf{U_2} \end{bmatrix} \qquad \mathbf{y} = 1 \tag{5.3}$$

The Big-M approach translates the same condition by assigning a large scaling value M that enforce the first condition if y=1 and the second condition if y=0 as follows:

$$\begin{aligned} \mathbf{z}-\mathbf{U}_1 \le \mathbf{M}(1-\mathbf{y}) \quad & \mathbf{z}-\mathbf{U}_2 \le \mathbf{M}\mathbf{y} \\ -\mathbf{z}+\mathbf{L}_1 \le \mathbf{M}(1-\mathbf{y}) \quad & -\mathbf{z}+\mathbf{L}_2 \le \mathbf{M}\mathbf{y} \end{aligned} \tag{5.4}$$

The disjunctive expression can be translated into the Big-M expression, and vice versa.

This study uses the Big-M expression for various logic conditions. Traditionally, a scalar value of **M** equal to the respective bounds $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{L}_1$ and $\mathbf{L}_2$ is selected to enforce constraint on a variable, as follows:

$$\begin{aligned} \mathbf{z}-\mathbf{y}\mathbf{U}_1 \le 0 \quad & \mathbf{z}-(1-\mathbf{y})\mathbf{U}_2 \le 0 \\ -\mathbf{z}+\mathbf{y}\mathbf{L}_1 \le 0 \quad & -\mathbf{z}+(1-\mathbf{y})\mathbf{L}_2 \le 0 \end{aligned} \tag{5.5}$$

Two or more equalities or inequalities can be combined together, into combined logic and algebraic expressions as follows:

$$\begin{aligned} \mathbf{y}(\mathbf{z}-\mathbf{U}_1)+(1-\mathbf{y})(\mathbf{z}-\mathbf{U}_2) \le 0 \\ \mathbf{y}(-\mathbf{z}+\mathbf{L}_1)+(1-\mathbf{y})(-\mathbf{z}-\mathbf{L}_2) \le 0 \end{aligned} \tag{5.6}$$

The above expressions show clearer logic relationships between the two operating conditions $\mathbf{L}_1 \le \mathbf{z} \le \mathbf{U}_1$ and $\mathbf{L}_2 \le \mathbf{z} \le \mathbf{U}_2$, and provides better translation of the disjunctive expression in (5.3). All expressions in (5.5) and (5.6) are used to model the MINLP superstructure problems in the case studies presented in section 5.5.

The common Big-M expressions in MINLP formulations are summarised from various sources (Bahri, 1996; Bemporad *et al.*, 1999; Raman and Grossmann, 1994) and presented in Table 5.1.

Table 5.1 Typical Big-M Expressions in MINLP Formulation

	Operational conditions and logic relationships	Mixed-integer (in)equalities
1	Selection among several a subset of units I	
	Select only one unit	$\sum_{i\in I} y_i = 1$
	Select at most one unit	$\sum_{i\in I} y_i \le 1$
	Select at least one unit	$\sum_{i\in I} y_i \ge 1$
2	If – then: If k exists j must exist	$y_k - y_j = 0$
3	Activating continuous variables: If $y = 1$ then $L_1 \le z \le U_1$ If $y = 0$ then $L_2 \le z \le U_2$	$z - U_1 \le M(1-y)$ $-z + L_1 \le M(1-y)$ $z - U_2 \le My$ $-z + L_2 \le My$
	Logic relationships	
4	And : Both k and j must exist	$y_k=1$ $y_j=1$
5	Or: At least one of k or j must exist	$y_j + y_k \ge 1$
6	Not: k must not exist	$1 - y_k$
7	Xor: Either one should exist, but not both	$y_j + y_k = 1$
8	Imply: If k exists, then j may exist	$y_k - y_j \le 0$
9	If and only if: If k exists, then j must exist	$y_k - y_j = 0$
10	Activating continuous variables: If $y = 1$ then $L_1 \le z \le U_1$ If $y = 0$ then $L_2 \le z \le U_2$	$(z - U_1)y + (z - U_2)(1-y) \le 0$ $(-z + L_1)y + (-z + L_2)(1-y) \le 0$

z = continuous variable
y = binary variable, $y = 1$.

Once the MINLP superstructure model has been developed, the next issue is to find the appropriate solution. Identifying the solution and application of MINLP problems is still an active research area, and the current available solutions are being continuously upgraded. The difficulties are due to the undifferentiability and potential exponential size of integer programming problem, as well as the integration between integer and continuous nonlinear programming.

An overview of MINLP algorithms and extensive theoretical, algorithmic, and applications-oriented descriptions of these algorithms can be found in Floudas (1995) and Grossman (2002). The approaches applied for chemical process synthesis include the Branch and Bound, the Generalised Benders Decompositions (GBD), and the Outer Approximation (OA) with additional variants, such as Equality Relaxation (OA/ER), and Augmented Penalty (OA/ER/AP). These approaches are briefly reviewed in section 5.2.4 as the background for the development of MINLP solver in section 5.4.1.

The additional complication to the MINLP process synthesis includes the solution of process dynamics, process nonlinearity and uncertainties, which give rise to dynamic, non-convex and semi-infinite optimisation problems, respectively. One of the major sources of these complications is the consideration of operability aspects in process synthesis, which is the aim of the proposed framework. The relevant works in this area include the consideration of flexibility and controllability in the original framework (Bahri, 1996), the interaction of design and control (Schweiger and Floudas, 1997), simultaneous economic and controllability studies (Heath *et al.*, 2000; Narraway and Perkins, 1994; Narraway *et al.*, 1991) and the stochastic approach on dynamic flexibility problems (Bansal *et al.*, 1998; Dimitriadis and Pistikopoulos, 1995; Mohideen *et al.*, 1996). While these works have been reviewed in Chapter 2, specific discussion on the respective algorithmic approaches to solve the dynamic, semi-infinite MINLP is given in section 5.2.2. This includes the comparison with the proposed framework.

5.2.3 MINLP Decompositions

The solutions to MINLP problems consist of solving three decomposition problems of the basic form of MINLP formulation in (5.1). The decomposition problems are the nonlinear programming (NLP) relaxation problems, the NLP subproblems for fixed **y**, the feasibility subproblems, and the mixed-integer linear programming (MILP) cutting plane subproblems. The utilisation of these subproblems determines the classifications of MINLP solution algorithms. Therefore, the subproblems are briefly described prior to the review of the algorithms.

5.2.3.1 NLP Relaxation

In the NLP relaxation problems, the discrete variables **y** are relaxed and considered to be continuous between lower and upper bounds. These bounds are the closest integer solutions, which are lower and higher than **y**, given **y** are non-integers. These bounds are updated and tightened iteratively during the MINLP search, until all discrete variables have integer solutions.

The objective value **Φ** yielded from the initial NLP relaxation is the absolute lower bound of the MINLP problem. The bounds of **y** are determined accordingly. These solutions may be sent to the successive NLP relaxation problems, or other computations. The successive NLP relaxation solutions are evaluated against the current bounds. If the new **Φ** is better than the current lower bound, and the new solutions of **y** are within the current respective bounds, all bounds are updated with the new values.

5.2.3.2 NLP Subproblems for Fixed Discrete Variables

The NLP subproblem is performed with all **y** fixed to certain integer solution. If it yields feasible solution, then the respective **Φ** becomes the new upper bound of the MINLP problem.

5.2.3.3 Feasibility Subproblem for Fixed Discrete Variables

This subproblem is performed if the corresponding NLP subproblem for fixed discrete variables does not obtain a feasible solution. It seeks for the infinity norm measurement of the respective infeasibility, and sends it back to the corresponding NLP subproblem. This is then used to generate the new relaxed **y**.

5.2.3.4 MILP Cutting Plane

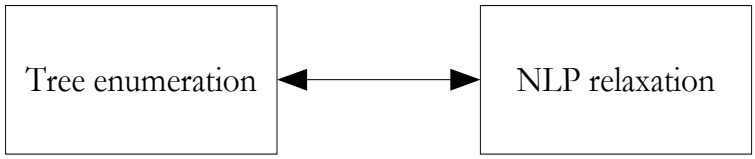
This subproblem provides the outer-approximation of the constraints, hence the feasible region, and the under-estimation of the objective function. The estimations are the linearisation of the constraints and objective functions at the solution of the preceding NLP subproblem. The linearisation may be performed fully on all constraints, or only on the violated subset. These linearised functions form a MILP problem, of which solution becomes the new lower bound of the MINLP problem.

5.2.4 MINLP Algorithms

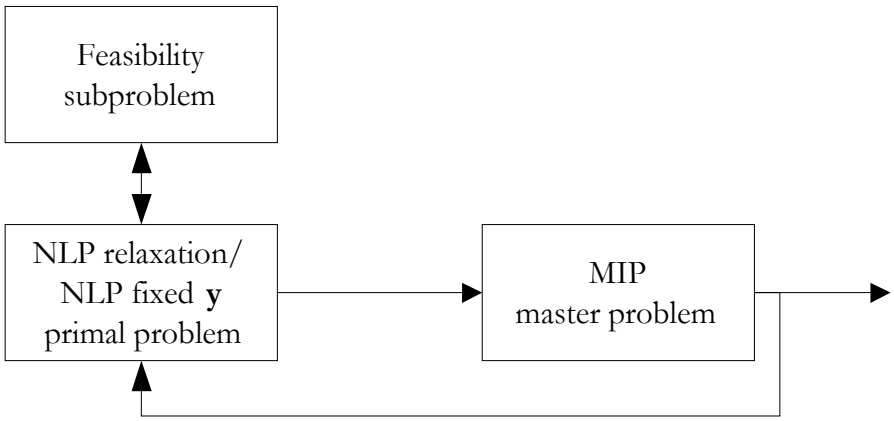
According to the utilisation of the subproblems, the following three major algorithms are reviewed (Grossman, 2002; Schweiger and Floudas, 1997):

1. Branch and Bound (BB)
2. Generalised Benders Decompositions (GBD)
3. Outer Approximation (OA) with additional features, such as Equality Relaxation (OA/ER), and Augmented Penalty (OA/ER/AP)

The organisation of the subproblems (Grossman, 2002) is illustrated in Figure 5.1, accompanying the brief review of each algorithm.



(a) Branch and Bound



(b) OA and GBD

Figure 5.1 The Structure of MINLP Algorithms

5.2.4.1 Branch and Bound

The Branch and Bound (BB) method is an intuitive approach to solve MINLP problems. It starts by solving the continuous NLP relaxation. If all **y** obtain integer values, then the search is stopped. Otherwise, a tree enumeration is performed in the space of **y**. The variables are successively fixed at the corresponding nodes of the tree. It gives rise to the relaxed NLP subproblem, which provides the new lower bound for the subproblems of the descendant nodes. The special case occurs when all **y** obtain integer solutions. In this case, the respective **Φ** becomes the new upper bound. A node is fathomed if the solutions are infeasible. The enumeration is completed when the yielded **Φ** exceeds the current upper bound.

5.2.4.2 Outer-Approximation

The Outer Approximation (OA) method (Duran and Grossmann, 1986; Fletcher and Leyffer, 1996) consists of performing a cycle of major iterations. It starts with primal NLP relaxation that yields the relaxed discrete solutions $\mathbf{y}^{k}$. Then, the MILP master problem is updated with the linearised constraints and objective function at the current solutions ($\mathbf{z}^{k},\mathbf{y}^{k}$). The corresponding solution includes the fixed discrete solutions $\mathbf{y}^{k+1}$, which are sent to the subsequent primal NLP subproblem. If the solution is feasible, the respective linearised functions are added to the MILP master problem. Otherwise, they are sent to the feasibility subproblem. The resulting infeasibility measurement is then used to get a new relaxed $\mathbf{y}^{k+1}$. The subsequent MILP master problem then yields a new fixed $\mathbf{y}^{k+1}$. The feasible fixed NLP subproblems yield the upper bounds of the MINLP problem. The cycle continues until the upper and lower bounds meet within a specified tolerance.

The OA algorithm requires some kind of integer-cut procedure to avoid returning to previous integer solutions $\mathbf{y}^{k}$ when convergence is achieved. The optimal solution of the MILP master problem is not required. Instead, it is sufficient to find a mixed integer solution ($\mathbf{z}^{k},\mathbf{y}^{k}$) within a given tolerance ε from the upper bound. In this case, the cycle is terminated when the MILP solution is infeasible.

5.2.4.3 Generalised Benders Decomposition

The Generalised Benders Decomposition (GBD) method (Geoffrion, 1972) is similar to the OA method. The difference is that the master problem is represented by a dual representation of the continuous space. The master problem only considers the violated inequalities and the discrete variables, while the continuous variables are optimised in NLP subproblems. Additionally, the initial primal problem is a NLP subproblem with fixed binary variables. (Floudas *et al.*, 1988).

It should be noted that the MILP master problems of both OA and GBD are commonly solved using the BB method.

5.2.4.4 Varian of Outer Approximation and Generalised Benders Decomposition Methods

Since OA and GBD rely on MILP solutions of the master problem, the consideration of nonlinear and nonconvex problems leads to several difficulties. Due to nonlinearity, the primal NLP problem sends different sets of locally linearised functions and solution points to master MILP problem in each cycle. This multiplies the size of the master problem during the cycle, and affects its convergence. The new appended set of linearised equalities may not be feasible when combined with the previously accumulated ones. Finally, if the problem is nonconvex, the resulting lower bounds of the master problem may miss out the global optimum.

To improve the convergence, Fletcher and Leyffer (1996) reformulate the master problem into a Mixed Integer Quadratic Programming (MIQP) problem, by using the Hessian of the Lagrangian of the preceding NLP solution. The accumulated number of constraints are reduced by aggregating the linear approximation of the objective and constraints (Grossman, 2002). Kocis and Grossmann (1987) address nonlinearity problems by relaxing the nonlinear equalities according to the sign of the Lagrange multiplier. This is referred to as the equality relaxation (OA/ER) variant. Viswanathan and Grossmann (1990) append slacks in the MILP master problem to reduce the likelihood of cutting-off feasible solutions, which is the further variation using the augmented penalties (OA/ER/AP). Furthermore, global convexity tests are applied to all linearisations with respect to the current solution (Grossmann and Kravanja, 1997).

Algorithmically, BB may delegate the nonlinearity and nonconvexity problems to the NLP subproblems. The efforts to improve the NLP capability in handling those problems include the generation of valid convex underestimators for the nonconvex functions. The generation of convex under estimators is based on a symbolic reformulation that transform nonconvex terms into simpler ones, such as bilinear, univariate concave, convex, linear fractional and simple power terms. These are achieved through the addition of new variables and constraints to the original NLP problem (Adjiman *et al.*, 1998a; b).

Further developments include the integration of BB algorithms with the NLP solutions. The α-BB algorithm generates the convex lower bounding function for twice differentiable nonconvex problems at each node (Adjiman *et al.*, 1998a; b). When convexity is not severe, Leyffer (2001) integrates the SQP method (for solving NLP subproblems) with BB, which enforces early detection and termination of infeasible or inferior NLP solutions.

5.2.4.5 The Development of MINLP Solutions

The available computer codes for solving MINLP problems are still limited. The program DICOPT (Viswanathan and Grossmann, 1990) is available in the modeling system GAMS (Brooke *et al.*, 1992) based on OA/ER/AP method. It is complemented with SBB as standard BB solver in GAMS (Brooke *et al.*, 1992). Leyffer (2001) implements the BB interlaced with the SQP algorithm in NLP subproblem, which is available in AMPL. Sahinidis (1996) implements BARON, the BB code with rigorous global optimisation capability for small convex problems. Schweiger and Floudas (1997) develop MINOPT, which implements the OA and GBD methods for solving dynamic MINLP problems.

The common challenges of these solutions include the formulation of MINLP problems, the development of preprocessor, branching or cut rules using information from any of the subproblems, and proper actions when an NLP subproblem fails. So far, the difficulties with nonconvexity and nonlinearity still force most MINLP solutions to accept local optimum solutions.

The OA codes strongly encourage the user to reformulate the MINLP as convex programming problem, keep the problem as linear and convex as possible, avoid the product of integer and continuous variables and remove the weak and redundant constraints. The problems due to nonconvexity in these codes are reported in the benchmark tests by Borchers and Mitchell (1996) and controllability studies by Perkins and co-workers (Heath *et al.*, 2000; Narraway and Perkins, 1994; Narraway *et al.*, 1991). They show that in nonconvex cases, OA iterates near the optimum solution without

ever finding the optimum. This problem is apparent when dynamic nonlinear problems is considered. It appears that the algorithms suffer from poor approximation of the nonconvex region. At some nonlinear points, the respective linearisation may also be numerically difficult and contribute further to the failures (Narraway and Perkins, 1994).

The master problem of GBD yields weaker lower bounds in comparison to OA, because it mainly considers discrete variables, which leads to longer iterations. While GBD is efficient when there are many discrete variables, often additional constraints must be added to the master problem to strengthen the bounds (Mohideen *et al.*, 1996). On the other hand, while OA has shorter cycles, the computational cost of solving the master problem is high because of the accumulated constraints during the iteration cycle (Grossman, 2002).

The BB codes are often associated with a potentially large tree enumeration of NLP subproblems, which may seem nonviable initially. Nevertheless, effective branching and bounding rules, problem formulations as well as improved NLP solutions can significantly reduced the number of nodes and improved the convergence. This include SBB in GAMS (Brooke *et al.*, 1992), and MINLP in AMPL (Leyffer, 2001). Only very recently, BB based MINLP solution has been available for use with MATLAB. It is available in the TOMLAB/MINLP package (Holmstrom *et al.*, 2003), which integrates the SQP algorithm in subproblems with BB algorithms (Leyffer, 2001). The solver is suitable for reasonably convex problems and quadratics objectives.

5.2.5 Operability Considerations in Process Synthesis

The operability considerations have extended the process synthesis problem to cover process dynamics (DAE solutions) and uncertainties. These require the integration of the MINLP solutions, with the dynamic and stochastic optimisation in some degrees. These solutions have been addressed separately in Grossman and Biegler (2003) and Sahinidis (2003). The focus in this section is on the combined approaches in process synthesis.

The early studies in this field include the consideration of flexibility in process synthesis and the design of new chemical processes (Saboo *et al.*, 1985), and the retrofit design of an existing process (Papalexandri and Pistikopoulos, 1994a; b; Pistikopoulos and Grossmann, 1988a; b). Other studies focus on the interaction between process design and process control (Nishida *et al.*, 1981).

The recent developments to address both flexibility and controllability include the original Dynamic Operability framework (Bahri, 1996), simultaneous economic and controllability studies (Heath *et al.*, 2000; Narraway and Perkins, 1994; Narraway *et al.*, 1991), the interaction of design and control (Schweiger and Floudas, 1997), and stochastic approach on dynamic flexibility problems (Bansal *et al.*, 2000; 2002a; b; Bansal *et al.*, 1998; Dimitriadis and Pistikopoulos, 1995; Mohideen *et al.*, 1996). The noticeable similarities amongst these studies are the usage of OA and GBD algorithms to solve the structure problems. The control structure selection, specifically the selection of suitable sets of manipulated and controlled variables and design of their interconnection, become prominent decision in pursuing the controllability objectives. The process control dynamics due to disturbances and uncertainties are solved in various degrees. It should be noted that due to nonlinear and nonconvexity of process dynamics, all of these applications guarantee local but not global optimal solutions.

Bahri (1996) solves the process synthesis problem of selecting the process configuration and deciding between two process control designs in a superstructure. The controllability is posed in terms of economic penalties that arise from the regulatory performances. In this case, the effecting external disturbances and uncertainties are characterised by a set of step functions with uniformly distributed magnitudes. The DAE are solved simultaneously with the NLP problem using the orthogonal collocation methods (Vassiliadis *et al.*, 1994). This method approximates the dynamic profiles with polynomials between collocation points, and includes the polynomial coefficients in the design variables.

The control parameters are pre-designed, then scaling factors are used to tune the parameters are further included in the design variables. The synthesis problem is

formulated as a large-scale, dynamic, semi-infinite MINLP problem, which is to be solved in two-level iterative Dynamic Operability Framework. The outer-level of the framework is the dynamic MINLP problem, which considers the critical disturbance and uncertainty combinations found in the inner level. The latter is an NLP of dynamic feasibility problem over the disturbance and uncertainty space, at a fixed design point given by the preceding outer level. If the process is found feasible in this level, the algorithm terminates. Otherwise, the critical disturbance and uncertainty combinations are identified and sent to the outer level. The dynamic MINLP problem in the outer level is solved using the OA/ER/PA code DICOPT in GAMS.

Perkins and co-workers (Heath *et al.*, 2000; Narraway and Perkins, 1994; Narraway *et al.*, 1991) concentrate on the control structure problem, which is the selection of measured variables and manipulated variables. They optimise the process economics, which are the penalties associated with the effect of disturbances to process output variables. The disturbances are characterised either as scalar values or step functions. The approach is restricted to stable processes, which are represented by index-1 DAE including time delay. In Narraway *et al.* (1991), the economy of linear process is computed by transforming the whole DAE, control and structure problem to the frequency domain, so that the dynamic state-space problem becomes a linear matrix algebra. The measurement of economic penalties, which is the maximum deviation from the active constraints in an open loop condition, is augmented to matrix computation and solved effectively. However, the consideration of dynamic nonlinear processes by solving the dynamics in the time domain within MINLP problem using OA/ER/PA code DICOPT, (Narraway and Perkins, 1994) fails to find the optimum solution due to nonconvexity. Heath (2000) addresses the problem by exploiting the OA/ER/PA structure. The dynamic optimisation problem is avoided by linearising the whole problem around the nominal steady state optimum point and transforming into the frequency domain. Hence, the MINLP problem is transformed into MILP. The enumeration size is reduced by enforcing the constraints of integral controllability with integrity (ICI); and the integral stabilisability (IS) (Campo and Morari, 1994) in the master problem.

Schweiger and Floudas (1998) directly compare the usage of OA and GBD code in solving dynamic MINLP problems with code MINOPT. Both approaches are demonstrated on multi-objective optimal control and synthesis problem of a chemical process affected by a single disturbance function. This review focuses on the required information at the master problems. While it is known that OA demands the binary variables to participate in a linear and separable fashion, Schweiger and Floudas find that the binary variables can not participate in the DAE. The reason is that the master problem does not handle the DAE directly, but through the respective linearised functions. These functions are considered as point constraints, and are derived using the gradient information at the solution of the primal problem. If the binary variables participate in the DAE, the yielded point constraints would consist of the original DAE, and therefore would not be linear and useful for master problem.

On the other hand, the master problem in GBD is formulated using the pure binary constraints and the Lagrange multiplier of all of the constraints including the DAE. The latter are dual variables, and those extracted from DAE are called adjoint variables. The associated adjoint problems are solved and used with other constraints to construct the master problem. It is claimed that the formulation has no restriction on whether or not binary variables participate in the DAE system. As long as the binary variables participate linearly, the master problem is always a MILP.

Mohideen *et al.* (1996) apply the GBD principles to optimal control structure selection of process with uncertainties in the two-level mixed integer dynamic optimisation (MIDO) framework. The uncertainties are discretised at the gaussian quadrature points over the parameter space. Then, the process is optimised at each point, giving rise to a multi-period feasibility optimisation problem, in which the respective objective of each problem is linearly combined using the associated load coefficients. The multi-period problem is in a fixed structure and regarded as the primal problem in GBD algorithm. The solution of primal problem is evaluated by a feasibility test over the disturbance and uncertainty space. The algorithm terminates if the process is feasible. Otherwise, the critical disturbance is extracted and added to the multi-period primal problem.

It is found that the dual information obtained from this problem is weak, due to participation of the binary variables in DAE. Therefore, an overall (adjoint) problem containing the active constraints of the primal solutions is solved to provide the required multiplier information. The master problem then provides a new set of control structures and a new lower bound for the next primal problem, subject to the accumulating Lagrangian functions and multiplier data generated from the overall problem.

Very recently, Bansal *et al.* (2000; 2002a; 2002b) introduced the flexibility analysis and synthesis using the parametric programming in the MIDO framework and revisited Mohideen's case. This framework obtains the parameters of the smallest set of linear equations, which tightly bound the feasible operational solution of the multi-period primal problem. These linear equations are sent to master problems, which therefore remove the requirement of explicit dual information and the adjoint problem of the primal solutions. It is also unaffected by the existence of binary variables in DAE.

The above cases show the obvious lack of BB algorithms in solving dynamic MINLP problem with uncertainties, except of course, to solve the respective master problems. This fact is worth further explorations, since BB guarantees finite enumeration nodes in comparison to OA and GBD. Given a tight branching rule and good formulations, the resulting convergence rate can be significantly increased. Furthermore, the nonlinearity and nonconvexity are mainly handled in NLP subproblems, not by BB algorithms. Therefore, as long as the quality of NLP solver is guaranteed, the optimality of MINLP solution may be stronger than OA and GBD.

In the next section, the development of a BB algorithm to solve the process synthesis problem in the proposed Dynamic Operability Framework is presented. The features added in this development include: the redundancy analysis and elimination arisen by each structure candidates, branching rules exploiting the binary constraints and binary list, combination of the depth-first and breadth-first search in tree enumeration, treatments for non-convergent and inferior NLP subproblems, and the capacity to change the priority of the binary decision.

The developed algorithm is parallel but independent to the Convex-hull Relaxation Problem (CRP) of the Generalised Disjunctive Programming (GDP) approach (Lee and Grossman, 2000). The GDP allows a combination of algebraic and logical equations to represents discrete decisions. The discrete decisions determining conditions in continuous space are represented with disjunctions, and purely discrete constraints are represented with logic propositions (Raman and Grossmann, 1994). The typical expression is the Conjunctive Normal Form (CNF) given in (5.2). The CRP reformulates the GDP as a tight MINLP of a convex-hull containing the whole constraints within the disjunctions (Vecchietti *et al.*, 2003). All continuous variables are considered as the sum of the disaggregated variables, which during the MINLP search converge to the respective condition within an optimum disjunctive. The number of constraints in CRP increases by $n_z+n_z\times n_y+1$, where n_z is the dimension of continuous variables and n_y is the number of disjunctions. Organised within the BB algorithm, CRP becomes the subproblem in the tree enumeration. Each subproblem typically converges to non-integer solutions of discrete variables associated with disjunctions. These non-integer values indicate their distance to all disjunctions. Therefore, the branching rule is to select the disjunction associated with highest non-integer values. While this GDP does not address DAE and uncertainty problems, it demonstrates faster convergence in comparison to the basic BB, OA and GBD with the Big-M formulations. It makes the GDP performance a good benchmark for our developed algorithm, which is demonstrated in section 5.5.1 and 5.5.2.

The development in the next section is more practical and empirical, in comparison with rigorous theoretical proofs contained in the respective reviewed studies. Nevertheless, the solution addresses the important issues in dynamic semi-infinite MINLP programming with satisfying performance.

5.3 Dynamic MINLP Formulation of DOF

Combining the formulations in Chapter 3 with the basic MINLP formulation in (5.1), the integration of the process and controller structure selection into the framework yields in the following formulation:

$$\min_{\bar{\mathbf{z}}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{5.7}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = \mathbf{0} \qquad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le \mathbf{0} \qquad \mathbf{j} \in \mathbf{I_r}$$

$$\mathbf{g_{jd}}(\mathbf{y}) \le \mathbf{0} \qquad \mathbf{jd} \in \mathbf{j}$$

$$\mathbf{r-OCI} = \frac{\max_{\boldsymbol{\theta}} \mathbf{AOS}_{\boldsymbol{\theta}}(\bar{\mathbf{w}}(t))_{t_0 \le t \le t_f} \cap \mathbf{DOS}(\bar{\mathbf{w}})}{\mathbf{DOS}(\bar{\mathbf{w}})} = \mathbf{1}$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{y}, \mathbf{x}_{ss}, \mathbf{u}_{ss}, \mathbf{w}_{ss}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z}^{\mathbf{l}} \le \mathbf{z} \le \mathbf{z}^{\mathbf{u}}\}$$

$$\mathbf{y} \in \{0,1\}$$

$$\boldsymbol{\theta}^{\mathbf{N}} \in \boldsymbol{\theta}^{\mathbf{k}} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\}$$

$$\bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}$$

$$\mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w}^{\mathbf{l}} \le \mathbf{w} \le \mathbf{w}^{\mathbf{u}}\}$$

$$\mathbf{x}_{ss} = \mathbf{x}(t = t_o)$$

$$\mathbf{u}_{ss} = \mathbf{u}(t = t_o)$$

$$\mathbf{w}_{ss} = \mathbf{w}(t = t_o)$$

$$t \in \{t : t_o \le t \le t_f\}$$

The nomenclature of this formulation is generally similar to the fixed structure version in Chapter 3. The difference is that in this formulation, $\bar{\mathbf{z}}$ contains the binary design variables **y** in addition to the existing continuous design variables **z**, $\mathbf{x_{ss}}$, $\mathbf{u_{ss}}$ and $\mathbf{w_{ss}}$. The functional variables $\bar{\mathbf{w}}$, which are used for **r-OCI** computation, are the subset of measured variables **w**. The process and controller dynamics are covered in $\mathbf{h_i}$, the steady state design constraints in $\mathbf{g_j}$, and dynamic constraints in **r-OCI**. The external disturbances as well as process uncertainties are characterised as a set of step functions with uniformly distributed magnitudes in **EDS**. Therefore, the above formulation is a semi-infinite dynamic mixed integer non-linear programming (SIDMINLP) problem. Similar to the fixed structure version, as well as the original approach (Bahri, 1996), the

semi-infinite programming problem in this framework is solved in an iterative, two-level dynamic optimisation algorithm as follows:

Outer-level:

$$\min_{\bar{\mathbf{z}}} \Phi(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{N}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t=\bar{t}} \tag{5.8}$$

$$\begin{aligned}
\text{s.t.}\quad & \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{\mathbf{k}}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 && \mathbf{i} \in \mathbf{E} \\
& \mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}^{k}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) \le 0 && \mathbf{j} \in \mathbf{I_r} \\
& \mathbf{g_{jd}}(\mathbf{y}) \le 0 && \mathbf{jd} \in \mathbf{j} \\
& \mathbf{r-OCI} = \frac{\mathbf{AOS_\theta}(\bar{\mathbf{w}}(t))_{t_0 \le t \le t_f} \cap \mathbf{DOS}(\bar{\mathbf{w}})}{\mathbf{DOS}(\bar{\mathbf{w}})} = 1 \\
& \bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{y}, \mathbf{x}_{ss}, \mathbf{u}_{ss}, \mathbf{w}_{ss}\} \\
& \mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z^l} \le \mathbf{z} \le \mathbf{z^u}\} \\
& \mathbf{y} \in \{0,1\} \\
& \boldsymbol{\theta}^{\mathbf{N}} \in \boldsymbol{\theta}^{\mathbf{k}} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\} \\
& \bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\} \\
& \mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w^l} \le \mathbf{w} \le \mathbf{w^u}\} \\
& \mathbf{x}_{ss} = \mathbf{x}(t = t_o) \\
& \mathbf{u}_{ss} = \mathbf{u}(t = t_o) \\
& \mathbf{w}_{ss} = \mathbf{w}(t = t_o) \\
& t = \{t : t_o \le t \le t_f\}
\end{aligned}$$

Inner-level:

$$\max_{\boldsymbol{\theta}} \mathbf{AOS_\theta}(\bar{\mathbf{w}}(t)) \tag{5.9}$$

$$\begin{aligned}
\text{s.t.}\quad & \mathbf{h_i}(\bar{\mathbf{z}}^*, \boldsymbol{\theta}(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t) = 0 \quad \mathbf{i} \in \mathbf{E} \\
& \bar{\mathbf{z}}^* \in \{\mathbf{z}^*, \mathbf{y}^*, \mathbf{x}_{ss}{}^*, \mathbf{u}_{ss}{}^*, \mathbf{w}_{ss}{}^*\} \\
& \bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\} \\
& \boldsymbol{\theta} \in \mathbf{EDS} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^{\mathbf{l}} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{\mathbf{u}}\} \\
& \mathbf{x_{ss}}^* = \mathbf{x}(t = t_{ow}) \\
& \mathbf{u_{ss}}^* = \mathbf{u}(t = t_{ow}) \\
& \mathbf{w_{ss}}^* = \mathbf{w}(t = t_{ow}) \\
& t \in (t : t_{ow} \le t \le t_f\}
\end{aligned}$$

In this formulation, $\mathbf{g}_{jd}$ is the subset of the steady state constraints $\mathbf{g}_j$ containing purely binary variables. This subset is used for the branching rule, which is discussed in section 5.4.3.

The major iteration within the framework is between the outer and inner levels. The outer level is a combined steady state and dynamic multi-period MINLP problem, which considers the regulatory process dynamics due to the critical disturbance and uncertainty combinations θ^k. The initial outer level only considers the nominal disturbances and uncertainties θ^N. The corresponding optimum solution $\bar{\mathbf{z}}^*$ found in the outer level is sent to the inner level, which performs the feasibility test for the framework. This level evaluates the process feasibility at $\bar{\mathbf{z}}^*$ and extracts the associated θ^k from the convex hull of AOS_θ. The latter is constructed from the dynamics of functional variables $\bar{\mathbf{w}}(t)$, which is found from the redundancy analysis and elimination procedure described in Chapter 4. The framework iteration completes if the outer level converges and the subsequent inner level does not introduce any additional θ^k. This typically happens in the second iteration.

In continuation of the case studies in Chapters 3 and 4, the outer level involves the nonlinearity and nonconvexity, which are inherent to the combined steady state and dynamic multi-period MINLP problem. Due to the well-known difficulties with these aspects in OA and GBD, as well as the lack of BB solutions, the solution to this problem is developed based on the BB algorithm. The development is described in the following section.

5.4 The Development of Dynamic MINLP Solution

The BB algorithm for the outer level of the proposed framework is developed based on the MATLAB functions *fminconset.m* (Solberg, 1999). This function operates on top of the existing MATLAB NLP solver *fmincon.m* (Mathworks, 2000), which is described

in the next section. Then, several issues, namely the superstructure MINLP model, the accommodation of redundancy analysis and elimination procedure, the branching rules, and the convergence of the NLP solutions are discussed for extension features of the algorithm. Finally, the issues are assembled in the proposed BB solution *fminconsete6.m*, and presented in section 5.3.6.

5.4.1 The Basic Algorithm

The *fminconset.m* function (Solberg, 1999) is a recursive BB algorithm that operates on top of the existing MATLAB NLP solver *fmincon.m* (Mathworks, 2000). It principally solves continuous optimisation problems, whilst constraining some variables into sets of standard values. These values may consist of discrete, integer or binary values. The associated discrete programming problem is recursively divided into two subproblems, by fixing the discrete variables to the closest above and below standard values.

The search starts by solving an NLP relaxation, and using the solution as the lower bound of the problem. If the solutions of the discrete variables are all equal to the values defined at the standard discrete set, then the optimum solution is reached and the search is stopped. Otherwise, the search branches on the first discrete variable that has non-standard solution. The closest discrete values above and below the current solution are identified. This is equivalent to the ceiling and floor functions of the current solution for integer problems discussed in section 5.2.4.1. If both above and below values exist, the NLP with the fixed above values becomes the first subproblem.

The first discrete variable with non-standard solution is identified. Subsequently, a new equality constraint to fix this variable to the above value is added to the original constraints, and the NLP subproblem subject to the updated constraints is solved. If the NLP subproblem converges, and yields the superior solution over the existing lower bound, then this solution becomes the new lower bound. The branching continues recursively to the next discrete value with non-standard solution. Otherwise, the node is fathomed. If this happens, the algorithm backtracks to the ascendant node, then resumes branching at the subproblem associated with below values.

This algorithm is a basic recursive, depth-first BB algorithm, which is simple and intuitive. It nevertheless may be extended to improve the convergence, especially to handle dynamic NLP subproblems. This algorithm and the proposed extension *fminconsete6.m* uses *fmincon.m* to solve the NLP subproblems, but they can also be used with other NLP solvers with some adjustments on passing variables. In the next four sections, the issues affecting the algorithm are discussed, leading to the assembly of *fminconsete6.m* in section 5.3.6.

5.4.2 Formulation of the Superstructure Problem

The *fminconset.m* does not enforce, or exploit any of the MINLP formulations. Nevertheless, the accuracy and convergence of the process synthesis solution is determined by the associated formulation of MINLP superstructure. The formulation involves the derivation of the process DAE from the mass and energy balances, characterisation of the uncertainties in some degrees of stochastic properties, and assignments of logic to every possible condition. These steps may require several iterations until a compact model is achieved.

This study considers disjunctive formulation as a good initial step to portray all possible conditions and logic relationships within a superstructure, and gradually combines and eliminates the logic expressions until the minimal combinations are achieved. Then, the disjunctive expressions are translated to binary expressions using guidelines in Table 5.1.

Several MINLP methods require additional decision variables, depending on the required internal solutions. To name a few, the simultaneous dynamic optimisation method (Cervantes and Biegler, 1997) uses the additional $n_w \times n_k$ decision variables, where n_w is the number of dynamic variables and n_k is the number of points used to approximate their dynamic profiles; the GDP requires additional $n_z+n_z \times n_y+1$, where n_z is the number of continuous variables in the disjunctions, and n_y is the number of disjunctions. These procedures multiply the problem size. As the DAE size grows, managing the variables becomes very complicated.

The proposed framework avoids those problems by solving the DAE sequentially within the NLP subproblems, and formulating the MINLP superstructure with the direct Big-M translations of the compact disjunctive formulations of the superstructure. This way, each discrete condition is represented by one binary variable, so does each dynamic variable. Therefore, the original number of decision variables and the problem size are maintained.

5.4.3 Branching Rules

Assuming minimum assignment of binary variables in the MINLP superstructure formulation, all combinations of the associated binary values become the convex-hull vertices of the disjunction problem. Unless redundant binary variables exist, enumerating the convex-hull vertices is equivalent to enumerating all possible conditions after all. Therefore, a stronger integer cut is required for reducing the enumeration size.

Given the current formulation, the NLP relaxation yields fractional solutions of binary variables. Suppose the binary variables are $\mathbf{y}_i$, i=1, 2, …, n_y. The *fminconset.m* selects the first node $y_1 = 1$ for the first branch, while keeping the remaining binary variables at their current values. The algorithm branches only if the NLP subproblem is convergent, feasible on all continuous and fixed binary variables, and superior to the current lower bound. This study changes these rules with five new ones.

The first rule is to select the node closest to the fractional solutions. Therefore, if the fractional solution of the first binary variable is less than 0.5, then the first main branch is $y_1 = 0$, otherwise it is $y_1 = 1$.

The second rule is to fix all other discrete variables to their respective closest binary values. This helps faster convergence to the binary solutions.

The third rule is to exploit the pure logic propositions within the MINLP formulations. Enforcing these constraints on branching rules would eliminate infeasible nodes from the tree. For example, if at least one of two sets of equipment

should operate, then the associated pure logic propositions would be $\mathbf{y_1}+\mathbf{y_2} \geq 1$. Suppose the NLP relaxation on this problem yields fractional solutions [0.4, 0.4], then the possible nodes in enumeration tree are $[y_1, y_2] \in [[0, 0], [0, 1], [1, 0], [1, 1]]$. Enforcing the constraint $\mathbf{y_1}+\mathbf{y_2} \geq 1$ prior solving the descendant NLP subproblem would eliminate, or fathom, the node [0, 0]. Based on the rules 1 and 2, the algorithm assesses [1, 1] first, then followed by [1, 0].

To accommodate evaluation of the feasible binary combinations that do not pass rules 1 and 2, for the above example [0, 1], the algorithm combines the depth-first and breadth-first searches instead of the pure depth-first search. This combined search evaluates the node [0, 1], after the assessment of [1, 1] and [1, 0]. It guarantees assessments of all feasible binary combinations starting from the strong ones. However, the breadth-first search may revisit the combinations previously checked at the depth-first search. This problem is prevented by the fourth rule.

The fourth rule is to check the feasible nodes against the list of previously evaluated binary combinations IL. This list is updated at each completed NLP subproblem, and passed along to all subsequent NLP subproblems. It contains all previously checked binary combinations, which are those converging to all binary solution, or on the other hand, those yielding both infeasible binary solutions and inferior objective value. Therefore, if a fixed binary combination passes rules 1, 2, and 3 in breadth-first search, but already listed in IL, then this combination is fathomed. The algorithm subsequently backtracks, and resumes the search at the other nodes.

The fifth rule is to decide whether to continue the search or to fathom the node based on the solutions at binary variables and the objective values Φ of NLP subproblems. The first NLP relaxation solution is used for the initial lower bound during the search.

The subsequent NLP subproblems may encounter the following conditions:

1. If the subsequent NLP subproblem converges, it yields all the binary solutions and the superior Φ, then the lower bound and IL are updated. The breadth-first search is then initiated for the remaining binary variables.

2. If the subsequent NLP subproblem converges, it yields superior Φ, but not all the binary solutions, then the lower bound is maintained. IL is updated with the initial binary inputs to the NLP subproblem, and the search is resumed at the binary variables with fractional solutions.
3. If the subsequent NLP subproblem does not converge, or does not yield superior Φ, then the node is fathomed and the current lower bound is maintained. The algorithm backtracks, and resumes the search at the other main branch.

The fifth rule is performed right after the completion of NLP subproblems, while the other four rules are performed within the breadth-search and depth-search, prior to solving the NLP subproblem. Once passing these first four rules, the associated process is subjected to redundancy analysis and elimination procedure.

5.4.4 Accommodation of the Redundancy Analysis and Elimination Procedure

The fixed binary combinations discussed in the previous section are associated with process and controller structure. They lead to the different dynamic characteristics, which are represented by functional variables $\overline{\mathbf{w}}$, and affected by critical disturbances θ^k, as discussed in Chapter 4. These functional variables and critical disturbances are required to generate the multi-period dynamic optimisation, and to compute the controllability index within the NLP subproblems. Therefore, once the binary combination passes the four branching rules, the redundancy analysis and elimination procedure can be performed on the associated problem. The resulting $\overline{\mathbf{w}}$ and θ^k are then passed to the NLP subproblems.

5.4.5 Treatments of NLP Convergence

Each NLP subproblem solves a dynamic, multi-period optimisation problem. For this purpose, this study uses *fmincon.m*, which applies the SQP method to solve medium

sized convex NLP problems. However, in general one may face nonlinearity and nonconvexity. These can cause severe difficulties on solving the NLP.

One of these difficulties is the infeasible quadratic problem (QP) within the SQP iteration, which leads to ill-conditioned approximation of Hessian of the Lagrangian of the associated problem. When this happens, the NLP 'loses' the direction to local optimum. If the NLP iterations are displayed, the user can observe this condition as a sudden jump in search direction and the infinite iterations, following an otherwise normal convergence.

The other problem, which is less serious, but time consuming, is slow convergence on inferior solutions. After several NLP iterations, it may show clear direction to inferior **Φ**, but may not immediately converge due to nonconvexity.

To handle these problems, two new conditions are added to the termination conditions of the NLP solver, as follows:

1. If the number of NLP iterations is greater than 3, the quadratic problem is infeasible, and the search direction is higher than 10, then the NLP subproblem is terminated.
2. If the number of NLP iterations yielding inferior **Φ** is greater than 10, then the NLP subproblem is terminated.

The numbers in those termination conditions are chosen heuristically, based on the average performance of *fmincon.m*. So far, these heuristics assist effective convergence without missing the known global optimum. For the proposed framework, the modification is performed in the functions *nlconste.m*, which is used by *fmincone.m* and *fminconsete5.m*, in place of the original MATLAB helper function *nlconst.m*.

This approach can also be applied to other NLP solvers, as long as the BB algorithm and NLP solver are independent of each other. Readers with interests to interdependency between BB and SQP are to refer the recent work by Leyffer (2001).

5.4.6 The Dynamic MINLP Algorithm

The discussed features are combined in the following algorithm:

Step 0: Set the lower bound of the MINLP $\Phi^* = \infty$, iteration counter k = 0, integer list IL = [ ], the objective function and constraints functions, initial conditions $\bar{\mathbf{z}}_0$ containing binary combination $\mathbf{y_0}$, functional variables $\bar{\mathbf{w}}^{\mathbf{k}}$ and critical disturbances $\boldsymbol{\theta}^{\mathbf{k}}$

Step 1: Perform NLP relaxation using the given objective function and constraints functions, $\bar{\mathbf{z}}_0$, $\bar{\mathbf{w}}^{\mathbf{k}}$ and $\boldsymbol{\theta}^{\mathbf{k}}$

Step 2: Evaluate NLP solutions Φ, $\bar{\mathbf{z}}^*$.

If (k=0 (first NLP relaxation) and $\Phi<\Phi^*$)

then (let $\Phi^*=\Phi$ and let k=index of binary variables **y**, and go to Step 3).

If (k≠0 (NLP subproblem) and $\Phi<\Phi^*$) and all **y** is binary

then (let $\Phi^*=\Phi$, update IL with **y*** (solution node fathomed),

let k=index of the remaining **y**, and return to the caller Step 5.a or 5.b).

If (k≠0 (NLP subproblem) and $\Phi<\Phi^*$ and not all **y** binary)

then (keep Φ^*, update IL with $\mathbf{y_0}$ (source node fathomed),

let k=index of non-binary solutions of **y**, and go to Step 3).

If (k=0 (first NLP relaxation) and NLP not convergent),

then (terminate, and back to step 0 with different $\bar{\mathbf{z}}_0$).

If (k≠0 (NLP subproblem) and ($\Phi\geq\Phi^*$, or NLP not convergent))

then (keep Φ^*, update IL with $\mathbf{y_0}$ (source node fathomed),

let k=index of non-binary solutions of **y**, and

return to the caller Step 5.a or 5.b).

Step3: Perform the combined depth-first and breadth-first search

If k exist, then (

if exist $\lceil y^*(k)$, then (let $y^a(k) = \lceil y^*(k)$, otherwise $y^a(k) = y^*(k)$),

if exist $\lfloor y^*(k)$, then (let $y^b(k) = \lfloor y^*(k)$, otherwise $y^b(k) = y^*(k)$),

if $|y^b(k)| > |y^a(k)|$, then (swap($y^b(k)$, $y^a(k)$)),

round the remaining of **y**, update**($\mathbf{y^b}$, $\mathbf{y^a}$)**, and

perform Steps 4-7), otherwise go to Step 8.

Step 4.a: If $\mathbf{y^a} \in \mathbf{IL}$, then (ignore $\mathbf{y^a}$, go to Step 4.b).

If $\mathbf{y^a} \notin \mathbf{IL}$ then (perform redundancy analysis to get $\overline{\mathbf{w}}^{\mathbf{k}}$ and $\boldsymbol{\theta}^{\mathbf{k}}$,

update the constraints with $y(k)=y^a(k)$)

Step 5.a: Go to Step 1 to perform NLP subproblem with $\mathbf{y^a}$, $\overline{\mathbf{w}}^{\mathbf{k}}$, $\boldsymbol{\theta}^{\mathbf{k}}$

and the updated constraints.

Wait for the returned $\mathbf{y_0^a}$.

Step 6.a: If (returned $\mathbf{y_0^a}$ does not exist or $\boldsymbol{\Phi}^a \geq \boldsymbol{\Phi}^*$) then ($\boldsymbol{\Phi}^a = \infty$ and go to Step 7)

If (returned $\mathbf{y_0^a}$ exists and $\boldsymbol{\Phi}^a < \boldsymbol{\Phi}^*$) then $\boldsymbol{\Phi}^* = \boldsymbol{\Phi}^a$

Step 4.b: If $\mathbf{y^b} \in \mathbf{IL}$, then (ignore $\mathbf{y^b}$, go to Step 7).

If $\mathbf{y^b} \notin \mathbf{IL}$, then (perform redundancy analysis to get $\overline{\mathbf{w}}^{\mathbf{k}}$ and $\boldsymbol{\theta}^{\mathbf{k}}$,

update the constraints with $y(k)=y^b(k)$

Step 5.b: Go to Step 1 to perform NLP subproblem with $\mathbf{y^b}$, $\overline{\mathbf{w}}^{\mathbf{k}}$, $\boldsymbol{\theta}^{\mathbf{k}}$

and the updated constraints.

Wait for the returned $\mathbf{y_0^b}$.

Step 6.b: If (returned $\mathbf{y_0^b}$ does not exist or $\Phi^b \geq \Phi^*$) then ($\Phi^b = \infty$ and go to Step 7)

If (returned $\mathbf{y_0^b}$ exists and $\Phi^b < \Phi^*$) then go to Step 7

Step 7: If there is no remaining k, then go to Step 8.

If the remaining k exists

If (returned $\mathbf{y_0^a}$ exist and $\Phi^a < \Phi^b$),

then (let $\Phi^* = \Phi^a$,

update process constraints with $y(k) = y^a(k)$,

update k = k+1, and go to Step 3).

If (returned $\mathbf{y_0^b}$ exist and $\Phi^b < \Phi^a$),

then (let $\Phi^* = \Phi^b$,

update process constraints with $y(k) = y^b(k)$,

update k = k+1, and go to Step 3).

Step 8: Stop

5.4.7 The Computational Complexity

The computational complexity of the above algorithm depends on four parameters. These are the number of feasible binary combination n_{yfeas} to solve using the BB algorithm, the number of iterations n_{iter} required to solve the corresponding NLP subproblem, the dimension of the corresponding functional variables n_w, and the time required to solve the DAE t_{DAE} within each subproblem.

The complexity subject to the functions of the above parameters is commonly stated using the 'big-O' notation $O(g_o(\mathbf{p}))$, where O stands for 'order' and represents the growth of a function of parameters $g_o(\mathbf{p})$. It is typical to say that the algorithm "takes time $O(p^3)$". This notation avoids stating specific constants of the functions and indicates only the dominant terms, highlighting that this function expresses the worst-

case behavior. For such functions, one is primarily interested in their rate of growth as **p** increases. An algorithm is said to have a polynomial time complexity if $g_o(\mathbf{p})$ is either a polynomial or a function bounded by a polynomial. Such algorithms include those of orders $O(\log\mathbf{p})<O(\mathbf{p})<O(\mathbf{p}\log\mathbf{p})<O(\mathbf{p}^2)<O(\mathbf{p}^3)<O(2^{\mathbf{p}})$. On the other hand, algorithms with complexities that cannot be bounded by polynomial functions are called exponential-time algorithms. These include the "exploding-growth" orders such as $O(\mathbf{p}!)$ (Nemhauser *et al.*, 1994; in Till *et al.*, 2003)

The complexity based on the number of feasible binary combinations n_{yfeas} relates to the size of search tree, which depends on the usage of the logic proposition $\mathbf{g_{jd}(y)}$ as described in section 5.4.3. At the best case, n_{yfeas} is linear to the number of binary variable n_y, while at the worst case it is equal to 2^{n_y}. Therefore, the complexity due to n_{feas} may vary between $O(n_y)$ to $O(2^{n_y})$. Since the algorithm requires recording the fathomed nodes in matrix **IL**, the maximum allowed size of **IL** becomes the limit of the possible combinations of n_y. In MATLAB, which is the software used in this study, the maximum allowed size of matrix is $2^{13}\times2^{13}$. Hence, the maximum n_y is 13.

The number of iterations n_{iter} depends on the linearity and the convexity of the problem. Obviously, the SQP solver in the proposed algorithm would solve linear and convex problems in a few iterations. At the worst case, where severe nonlinearity and nonconvexity occurs, n_{iter} may reach the maximum number of iteration $n_{maxiter}$ defined for the solver. The possibility has been reduced with the condition given in section 5.4.6 step 2, and the treatment of NLP convergence in section 5.4.5. These conditions require that each NLP subproblem converge before $n_{maxiter}$ is reached, and limits the number of iteration of inferior NLP subproblem to 10. At either cases, the complexity is linear to n_{iter}, hence $O(n_{iter})$.

The dimension of functional variables n_w determines the ease of geometric computation of AOS_θ, DOS and r-OCI within the NLP subproblems. At the current stage, the corresponding *Qhull* can handle up to 6-dimensional polyhedron with large

number of points within the hull in a reasonable speed (Barber, 1996). The computation grows in order $O(n_w!)$, hence gives the tractability limit of the problem.

The time required for solving the DAE, t_{DAE}, depends on the DAE index i_{DAE} and the total numbers of state and algebraic variables n_{xw}. The utilised DAE solver naturally handles DAE index 1; and as explained in chapter 3, the algorithm structure empirically enables handling DAE index-2. The latter gives the tractability limit of the problem, while the complexity grows linearly with n_{xw}; hence gives $O(i_{DAE} \times n_{xw})$.

Overall, the worst case complexity of the algorithm that limit the tractability of the problem is $O(2^{n_y} \times n_{iter} \times n_w! \times i_{DAE} \times n_{xw})$. Memory-wise, the tractability is limited by n_y and n_w, while the computational time is determined by n_{iter}, i_{DAE} and n_{xw}. as described earlier, the maximum numbers of n_y, n_w, i_{DAE} and n_{iter} are 13, 6, 2 and $n_{maxiter}$ respectively. The maximum numbers of n_{xw}, however is still an open problem. Recently, Till *et al.* (2003) reported detailed discussions about the complexity of MILP problems. A comparable analysis on dynamic MINLP cases would provide answer to the above open problem, therefore it is recommended for future work.

5.4.8 The Implemented Function

The algorithm is implemented in *fminconsete6.m* with some modification of the data files and passing variables. The inputs are as follows: fun, $\bar{\mathbf{z}}_0$, intidx, intlist, **A**, **B**, **Aeq**, **beq**, **lb**, **ub,** nonlcon$_1$, nonlcon$_2$, options, set, J_{min}, k_2, **ED**S, $\boldsymbol{\theta}^k$, cp, p_1, p_2. The outputs are: $\bar{\mathbf{z}}^*$**, Φ,** exitflag, output, λ, δ, hessian, intlist. These inputs and outputs are assumed as matrices. The unavailable information is supplied as empty matrix [ ]. The descriptions of the inputs and the outputs are given in Tables 5.2 and 5.3. The shaded parts in the tables are the new functions and passing variables used in *fminconsete6.m*, while the rests are originally used by *fminconset.m and fmincon.m*.

Table 5.2 The Inputs of *fminconsete6.m*

Input	Descriptions
fun	Function calculating the objective function Φ at $\bar{\mathbf{z}}$ The function accepts $\bar{\mathbf{z}}$ and returns the scalar Φ
$\bar{\mathbf{z}}_0$	An array of initial values of $\bar{\mathbf{z}}$
intidx	The indices of binary variables **y** as subset of $\bar{\mathbf{z}}$
intlist	List of fathomed binary solution. Default = [ ]
A, **B**	Matrices representing linear inequalities $\mathbf{A}\bar{\mathbf{z}} \leq \mathbf{B}$
Aeq, Beq	Matrices representing linear inequalities $\mathbf{Aeq}\,\bar{\mathbf{z}} \leq \mathbf{Beq}$
Lb, Ub	A set of lower and upper bounds of $\bar{\mathbf{z}}$, $\mathbf{Lb} \leq \bar{\mathbf{z}} \leq \mathbf{Ub}$
Nonlcon_1	Function computing nonlinear constraints of $\bar{\mathbf{z}}$. The function nonlcon1 accepts $\bar{\mathbf{z}}$ and returns the vectors **C** and **Ceq**, representing the nonlinear inequalities and equalities respectively.
Nonlcon_2	Function computing pure binary constraints of **y** and performing redundancy analysis and elimination of $\bar{\mathbf{z}}$. The function nonlcon2 accepts $\bar{\mathbf{z}}$ and **EDS** and returns the matrices $\boldsymbol{\theta}^k$ and **cp**.
options	Default optimisation parameters
set	Cell arrays of allowed binary values for each entry in $\bar{\mathbf{z}}$. Empty cell if the entry is continuous. Example: for $\bar{\mathbf{z}}(1) = 0$ or 1, and $\bar{\mathbf{z}}(2)$ continuous, then set = ([0 1],[ ])
J_{min}	The lower bound of MINLP solution so far. Default = Inf.
k_2	The index of binary variables evaluated at the current node
EDS	All disturbance and uncertainty combinations
θ^k	Critical disturbances
cp	Index of functional variables $\bar{\mathbf{w}}^{\mathbf{k}}$
$p_1, p_2, \ldots$	Problem dependent parameters $p_1, p_2, \ldots$

Table 5.3 The Outputs of *fminconsete6.m*

Input	Descriptions
$\bar{\mathbf{z}}^*$	Optimal value of $\bar{\mathbf{z}}$
$\boldsymbol{\Phi}$	Optimum objective value at the solution $\bar{\mathbf{z}}$
exitflag	Exit condition of the NLP subproblem If exitflag > 0 then fmincon NLP subproblem converges to a solution $\bar{\mathbf{z}}^*$ If exitflag $= 0$ then maximum number of iteration is reached If exitflag < 0 then NLP subproblem does not converge
output	Output properties information, such as number of iterations, methods, etc.
λ	Lagrange multipliers at the solution $\bar{\mathbf{z}}^*$
δ	Gradient of $\boldsymbol{\Phi}$ at the solution $\bar{\mathbf{z}}^*$
Hessian	Hessian of $\boldsymbol{\Phi}$ at the solution $\bar{\mathbf{z}}^*$
intlist	List of fathomed binary combination so far

The features of this function in comparison to the basic algorithm are as follows:

1. The original *fmincon.m* and *fminconset.m* require two input files. The first file 'fun' contains the computation of the objective value $\boldsymbol{\Phi}$, and if available, its gradient and Hessian. The second file 'nonlcon$_1$' contains the nonlinear equalities and inequalities $\mathbf{h}_i$ and $\mathbf{g}_j$; as well as their gradients and hessians, if available. The proposed *fminconsete6.m*, requires an additional file 'nonlcon$_2$', which contains the integer feasibility test, and the redundancy analysis and elimination procedure.
2. The indices 'intidx' indicate the position of $\mathbf{y}$ as subset of $\bar{\mathbf{z}}$. They can be ordered to represent their priority, from the most important to the least.
3. The lower bound is determined by the first NLP relaxation solution. The solution of the subsequent subproblem updates the lower bound only if it has all binary solutions, and if it is superior from previously found binary solutions.

The performance of this algorithm is demonstrated in the next case studies.

5.5 Case Studies

This section presents four case studies of process and controller synthesis using the proposed dynamic MINLP solution within the Dynamic Operability Framework. The aims of these case studies are as follows:

1. Demonstration of the MINLP formulations of the given superstructures to accommodate process and controller structure selection.
2. Demonstration of the tree enumeration and the embedded redundancy analysis and elimination performances.

The case studies consist of two steady state and convex network superstructure cases and two dynamic, semi-infinite, nonlinear CSTR superstructure cases. The first two cases do not address the dynamic and semi-infinite nature of the proposed solution, but focus on the alternative logic assignments within the MINLP formulations and the benchmark of convergence against the known solutions given in the open literature. The other two cases are the extensions of CSTR case studies presented in Chapters 3 and 4. Both cases address the existence of binary variables in DAE, which is traditionally not to be solved with OA approach (Schweiger and Floudas, 1998). The first CSTR case demonstrates the control structure selection on a fixed process structure, providing the continuation of the previous CSTR cases. The second CSTR case demonstrates both control and process structure selection problem, providing interesting illustration of interaction between process and control structure in achieving process controllability. These case studies are computed using Pentium IV 2.4 GHz Personal Computer with 256 MB RAM.

5.5.1 Network Superstructure 1

5.5.1.1 Problem Formulation

The first network superstructure is a disjunctive problem taken from Grossmann and Kravanja (1997) and Lee and Grossman (2000), which is illustrated in Figure 5.2. The network consists of three units $y_{p1\text{-}3}$, which process two streams $w_{1\text{-}2}$.

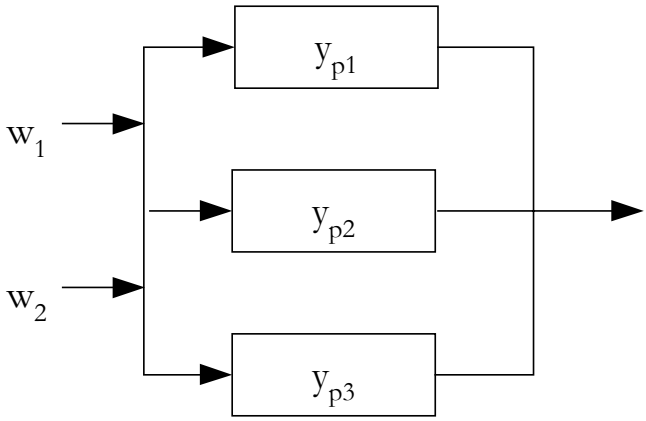


Figure 5.2 Network Superstructure 1: Process Structure Selection

The problem is to find the combination that minimise the cost **Φ**:

$$\Phi = y_{p1} + 1.5y_{p2} + 0.5y_{p3} + w_1^2 + w_2^2 \tag{5.10}$$

Subject to the following constraints:

$$\begin{aligned}
&g_1 : (w_1 - 2)^2 - w_2 \le 0 && g_7 : (w_1 - w_2) - 4y_{p3} \le 0 \\
&g_2 : w_1 - 2y_{p1} \ge 0 && g_8 : (w_1 + w_2) - 3y_{p3} \ge 0 \\
&g_3 : (w_1 - w_2) - 4y_{p1} \le 0 && g_9 : w_1 - y_{p3} \ge 0 \\
&g_4 : w_1 - w_2 y_{p2} \le 0 && g_{10} : y_{p1} + y_{p2} + y_{p3} \ge 1 \\
&g_5 : w_1 - y_{p2} \ge 0 && g_{11} : w_1 \le 4 \\
&g_6 : w_2 - y_{p2} \ge 0 && g_{12} : w_2 \le 4 \\
& && g_{13} : 0 \le w_1 \\
& && g_{14} : 0 \le w_2
\end{aligned} \tag{5.11}$$

In the set (5.11), constraints g_1 and g_{11}-g_{14} specify the amount of w_1 and w_2 in all unit processes, while g_2-g_9 assign the disjunctive operation for each unit. Constraint g_{10} states that at least one processing unit should exist, and used as the branching rule. The integer assignments in g_2-g_9 exploit the Big-M approach given in Table 5.1 to enforce the respective constraints.

5.5.1.2 Framework Implementation

Since this case only considers steady state process with no indication of disturbances or uncertainties, the framework is only performed at the first outer level. It focuses on evaluating the MINLP framework performance in selecting the process structure, especially with the given logical process assignments in the set (5.11).

The MINLP solver starts with NLP relaxation problem, which initiates the tree enumeration of the binary variables. If this subproblem converges to all-integer solutions for $y_{p1\text{-}3}$, the cost becomes the first lower bound of the optimal solution, the node is fathomed, and the other nodes are checked for superior solutions. If any of $y_{p1\text{-}3}$ has integer solution, then the node branches at these particular variables. The enumeration starts with the branching rule given in section 5.4.3, and continues until all feasible nodes are accounted for.

5.5.1.3 Results and Discussions

The MINLP solution converges in three nodes and 0.33 CPU seconds regardless of the initial conditions. The results and the tree enumeration are shown in Tables 5.4-5.5. In Table 5.5, node 0 indicates the NLP relaxation subproblem, which sets the lower bound Φ^*. Since y_{p1} and y_{p2} found in this subproblem are not integer, for the second node they are fixed to the closest binary values, 0 and 1, respectively. This node finds the integer solutions of y_{p1} and y_{c2}, and therefore updates the Φ^*. The third node gives inferior solution, hence fathomed. Since there are no other feasible binary combinations, the search is terminated and node 1 becomes the optimum solution. This result complies with the one found by Lee and Grossman (2000).

Table 5.4 Superstructure 1: Synthesis Results

	Steady state optimum
Φ	3.5
W_1	1
W_2	1
y_{p1}	0
y_{p2}	1
y_{p3}	0

Table 5.5 Superstructure 1: Tree Enumeration

Node no		0	1	2
Initial values	y_{p1}	0	0	1
	y_{p2}	0	1	1
	y_{p3}	0	0	0
Generated results	y_{p1}	0.1182	0	1
	y_{p2}	0.8818	1	0.6355
	y_{p3}	0	0	0
	Φ	3.479	**3.5**	**6.1768**

Optimum solution Fathomed

5.5.2 Network Superstructure 2

5.5.2.1 Problem Formulation

This case study is the network superstructure also taken from Lee and Grossman (2000), and illustrated in Figure 5.3.

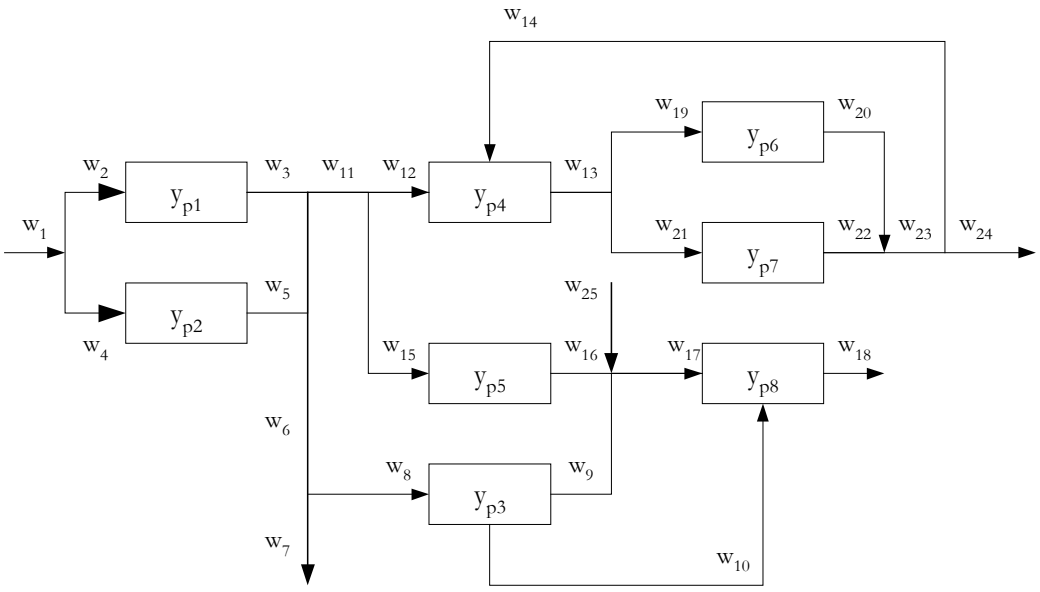


Figure 5.3 Network Superstructure 2: Process Structure Selection

The network consists of eight unit processes $y_{p1\text{-}8}$ which each has different capacity of converting the streams $w_{1\text{-}25}$. The problem is to find the combination that minimises the cost Φ defined as follows:

$$\Phi = c^T y_p + a^T w + 122 \tag{5.12}$$

$c^T =$ [5 8 6 10 6 7 4 5]

$a^T =$ [0 1 -10 1 -15 0 0 0 -40 15 0 0 0 15 0 0 80 -65 25 -60 35 ...
-80 0 0 -35]

Subject to the following constraints:

$$g_1 : w_1 - w_2 - w_4 = 0 \qquad g_7 : w_6 - w_7 - w_8 = 0 \tag{5.13}$$
$$g_2 : w_{13} - w_{19} - w_{21} = 0 \qquad g_8 : w_3 + w_5 - w_6 - w_{11} = 0$$
$$g_3 : w_{17} - w_9 - w_{16} - w_{25} = 0 \qquad g_9 : w_{11} - w_{12} - w_{15} = 0$$
$$g_4 : w_{23} - w_{20} - w_{22} = 0 \qquad g_{10} : w_{23} - w_{14} - w_{24} = 0$$

$$g_5 : w_{10} - 0.8w_{17} \le 0 \qquad g_{11} : w_{10} - 0.4w_{17} \ge 0$$
$$g_6 : w_{12} - 5w_{14} \le 0 \qquad g_{12} : w_{12} - 2w_{14} \ge 0$$

$$g_{13} : (\exp(w_3) - 1 - w_2)y_{p1} + (w_3 + w_2)(1 - y_{p1}) = 0 \tag{5.14}$$
$$g_{14} : (\exp(w_5 / 1.2) - 1 - w_4)y_{p2} + (w_4 + w_5)(1 - y_{p2}) = 0$$
$$g_{15} : (1.5w_9 - w_8 + w_{10})y_{p3} + (w_9 + (w_8 - w_{10}))(1 - y_{p3}) = 0$$
$$g_{16} : (1.25(w_{12} + w_{14}) - w_{13})y_{p4} + (w_{12} + w_{13} + w_{14})(1 - y_{p4}) = 0$$
$$g_{17} : (w_{15} + 2w_{16})y_{p5} + (w_{15} + w_{16})(1 - y_{p5}) = 0$$
$$g_{18} : (\exp(w_{20}) / 1.5 - 1 - w_{19})y_{p6} + (w_{19} + w_{20})(1 - y_{p6}) = 0$$
$$g_{19} : (\exp(w_{22}) - 1 - w_{21})y_{p7} + (w_{21} + w_{22})(1 - y_{p7}) = 0$$
$$g_{20} : (\exp(w_{18}) - 1 - w_{10})y_{p8} + (w_{10} + w_{17} + w_{18})(1 - y_{p8}) = 0$$

$$g_{21} : y_{p1} - (y_{p3} + y_{p4} + y_{p5}) \le 0 \tag{5.15}$$
$$g_{22} : y_{p2} - (y_{p3} + y_{p4} + y_{p5}) \le 0$$
$$g_{23} : y_{p3} - (y_{p1} + y_{p2}) \le 0$$
$$g_{24} : y_{p4} - (y_{p1} + y_{p2}) \le 0$$
$$g_{25} : y_{p5} - (y_{p1} + y_{p2}) \le 0$$
$$g_{26} : y_{p3} - y_{p8} \le 0$$
$$g_{27} : y_{p5} - y_{p8} \le 0$$
$$g_{28} : y_{p4} - (y_{p6} + y_{p7}) \le 0$$
$$g_{29} : y_{p6} - y_{p4} \le 0$$
$$g_{30} : y_{p7} - y_{p4} \le 0$$
$$g_{31} : y_{p8} - y_{p3} + y_{p5} + (1 - y_{p3})(1 - y_{p5}) \le 0$$

$$g_{32} : y_{p1} + y_{p2} = 0 \tag{5.16}$$
$$g_{33} : y_{p4} + y_{p5} = 0$$
$$g_{34} : y_{p6} + y_{p7} = 0$$

In the set (5.13), the constraints $g_{1\text{-}12}$ are the material balances, g_{13}-g_{20} represent the logical operations of the units, and g_{21}-g_{34} are the Big-M translation of logic propositions within the process. The constraints used for branching rules are g_{21}-g_{31}.

5.5.2.2 Framework Implementation

Similar to the previous network case study, the framework is performed at the first outer level. This time, the MINLP performance is evaluated on handling tight and combined disjunctive operation g_{13}-g_{20}, as well as more logic propositions g_{21}-g_{34}.

5.5.2.3 Results and Discussions

The MINLP solution converges in six nodes, if all initial conditions are zero; and three nodes, if all of them are higher than 0.4. The optimum solutions are the same in both cases, and comply with that given in Lee and Grossman (2000). The synthesis result and the three nodes tree enumeration are shown in Tables 5.6 and 5.7. This computation is completed in 2.7 CPU seconds.

Table 5.6 Superstructure 1: Synthesis Results

	Steady state optimum
Φ	64.2366
y_1	0
y_2	1
y_3	0
y_4	1
y_5	0
y_6	1
y_7	0
y_8	1

Table 5.7 Superstructure 1: Tree Enumeration

Node no		0	1	2
Initial values	y_1	0.5	0	0
	y_2	0.5	1	1
	y_3	0.5	0	0
	y_4	0.5	1	1
	y_5	0.5	1	0
	y_6	0.5	1	0
	y_7	0.5	1	0
	y_8	0.5	0	0
Generated results	y_1	0	0	0
	y_2	1	1	1
	y_3	0	0	0
	y_4	0.5	1	1
	y_5	0.5	0	0
	y_6	0.5	1	1
	y_7	0.5	0	0
	y_8	1	1	0
	Φ	106.8789	**64.2366**	**75.1534**

Optimum solution Fathomed

5.5.3 Controller Structure Selection

5.5.3.1 Problem Formulation

The first case study is the controller structure selection of the CSTR superstructure given in Figure 5.4. The basic description of the process has been given in Chapter 3. In this case study, the process is extended into a superstructure to accommodate the controller structure problem, which is used in deciding the optimum pairing between measurement and manipulated variables to control each reactor. The possible measurements are the product temperatures T_1, T_2, and the cooling amounts $Cool_1$, $Cool_2$. For any selected measurements, the manipulated variables are the coolant flow rates m_{c1} and m_{c2}. All possible control structures are assessed subject to the expected disturbances, which are the feed temperature T_f and the feed composition C_f.

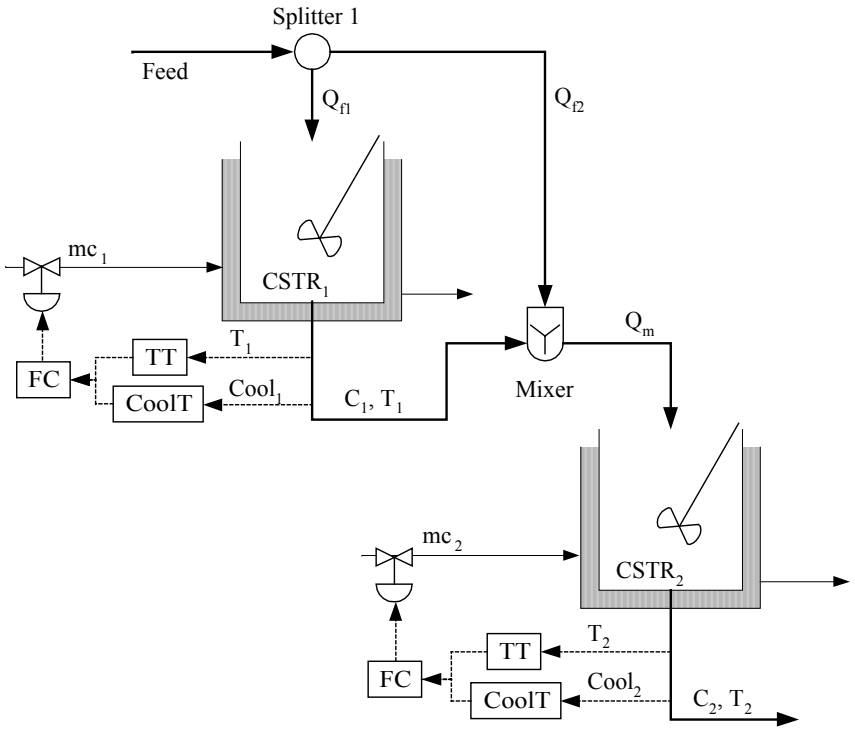


Figure 5.4 Superstructure 1: Controller Structure Selection

The control equations of this superstructure are as follows:

$$e_{11} = y_{c1}\left(T_{1ss} - T_1\right) \tag{5.17}$$

$$\frac{dI_{11}}{dt} = e_{11} \tag{5.18}$$

$$e_{12} = \left(1 - y_{c1}\right)\left(Cool_{1ss} - Cool_1\right) \tag{5.19}$$

$$\frac{dI_{12}}{dt} = e_{12} \tag{5.20}$$

$$m_{c1} = m_{c1ss} + \alpha_1\left(K_{c11}\left(e_{11} + \frac{I_{11}}{\tau_{11}}\right) + K_{c12}\left(e_{12} + \frac{I_{12}}{\tau_{12}}\right)\right) \tag{5.21}$$

$$e_{21} = y_{c2}\left(T_{2ss} - T_2\right) \tag{5.22}$$

$$\frac{dI_{21}}{dt} = e_{21} \tag{5.23}$$

$$e_{22} = \left(1 - y_{c2}\right)\left(Cool_{2ss} - Cool_2\right) \tag{5.24}$$

$$\frac{dI_{22}}{dt} = e_{22} \tag{5.25}$$

$$m_{c2} = m_{c2ss} + \alpha_2\left(K_{c21}\left(e_{21} + \frac{I_{21}}{\tau_{21}}\right) + K_{c22}\left(e_{22} + \frac{I_{22}}{\tau_{22}}\right)\right) \tag{5.26}$$

Two binary variables, y_{c1} and $y_{c2,}$ are assigned to facilitate the selection between T and Cool as measurement variable on each reactor. The variables are assigned directly at the measurement candidates in (5.17), (5.19), (5.23), and (5.25).

To reduce the complication of control tuning during the optimisation, the associated PI controllers are pre-designed using the Ziegler-Nichols method (Bahri, 1996; Ziegler and Nichols, 1943). Then, the optimum controller parameters are tuned further using the scaling factors α_1 and α_2. The logic assignments and associated controller parameters are given in Table 5.8.

Table 5.8. Controller Parameters and Binary Assignments

Control loop		Binary	Proportional	Reset
idx	Control pair	Assignment	gain K_i	times τ_i
11	T_1 - m_{c1}	$y_{c1} = 1$	-0.0023	0.1667
12	$Cool_1$ - m_{c1}.	$y_{c1} = 0$	0.001554	0.1668
21	T_2 - m_{c2}	$y_{c2} = 1$	-0.0023	0.1667
22	$Cool_2$ - m_{c2}.	$y_{c2} = 0$	0.01456	0.1668

idx = controller parameter index

This process is optimised to maximise the following nominal net profit Φ_N.

$$\Phi_N = 10(Q_{f1}C_{fN} + Q_{f2}C_{fN} - 0.3(Q_{f1} + Q_{f2}))\ldots \\ -0.01Cool_1 - 0.1Cool_2 - 0.1(Q_{f1} + Q_{f2}) \tag{5.27}$$

This time, the process operational constraints include four additional control tuning specifications $g_{9\text{-}12}$ in the following set:

$$\begin{aligned} &g_1 : Q_{f1} \geq 0.05 && g_2 : Q_{f2} \geq 0.05 \\ &g_3 : Q_{f1} + Q_{f2} \leq 0.8 && g_4 : C_2 \leq 0.3 \\ &g_5 : T_1 \leq 350 && g_6 : T_2 \leq 350 \\ &g_7 : Cool_1 \leq 30 && g_8 : Cool_2 \leq 20 \\ &g_9 : \alpha_1 \geq 0 && g_{10} : \alpha_2 \geq 0 \\ &g_{11} : \alpha_1 \leq 2 && g_{12} : \alpha_2 \leq 2 \end{aligned} \tag{5.28}$$

The arrangements in Equations (5.17), (5.19), (5.23), and (5.25), Table 5.8 and Constraints (5.28) cover all three fixed controller structure case studies in Chapter 3, as well as other possible structures. For instance, the open-loop structure is given when all α=0; the single-PI structure is when α_1=1, y_{c1}=1, and α_2=0; and the double-PI structure controlling T_1 and $Cool_2$ is when α_1=α_2=1, y_{c1}=1, and y_{c2}=0. Other possible structures include the double-PI control of both T_1 and T_2 (y_{c1}=y_{c2}=1, α_1=α_2=1), and the control of both $Cool_1$ and $Cool_2$ (y_{c1}=y_{c2}=1, α_1=α_2=1). Since there is no logic proposition on y_{c1} and y_{c2} in (5.28), all structures have equal probabilities.

5.5.3.2 Framework Implementation

Within the proposed framework, the models, the objective function, and the constraints are arranged according to the formulation (5.8)-(5.9). The initial outer-level contains only the steady state model, which consists of the steady state version of (3.17)-(3.25) plus the profit equation (5.10). The inner-levels solve the DAE model, which consists of the dynamic version of DAE (3.17)-(3.25) plus the control and profit equations (5.17)-(5.27). The critical disturbance combinations found in the inner-levels are sent to the subsequent outer-levels. These particular outer-levels seek for the optimal operating conditions subject to the critical disturbance combinations by solving both steady state and DAE models. The DAE outputs are $Cool_1$, $Cool_2$, m_{c1}, m_{c2}, and Φ_N; in addition to the state variables T_1, T_2, C_1, C_2, I_{11}, I_{12}, I_{21}, and I_{22}. The DAE are simulated over 100 seconds time horizon with 1-second intervals.

The framework seeks for the optimum values of design variables Q_{f1}, Q_{f2}, α_1, α_2, y_{c1}, y_{c2} as well as the initial conditions of T_{1ss}, T_{2ss}, C_{1ss}, C_{2ss}, $Cool_{1ss}$, and $Cool_{2ss}$. These variables are constrained by (5.28), in accordance with $\mathbf{g_j}(\overline{\mathbf{z}}, \boldsymbol{\theta}^k, \mathbf{x}, \dot{\mathbf{x}}, \mathbf{u}, \mathbf{w}, \mathbf{p}, t) \le 0$, $\mathbf{j} \in \mathbf{I_r}$ in (5.8). Furthermore, the constraints $\mathbf{g_{4\text{-}8}}$ also apply to the initial conditions as well as the whole dynamic profiles of T_1, T_2, $Cool_1$, $Cool_2$ and C_2, which determine the process controllability index **r-OCI**. The associated DOS and AOS_θ are constructed using the subset of these variables, which are selected using the embedded redundancy analysis and elimination procedure (see section 5.4.4). The binary variables y_{c1} and y_{c2} give rise to a four nodes enumeration tree of the dynamic MINLP problem. Each node of the tree represents one control structure, as well as the associated functional variables $\overline{\mathbf{w}}^k$ and critical disturbance combinations $\boldsymbol{\theta}^k$.

Since the first outer-level only considers steady state process and nominal disturbances, it does not perform any redundancy analysis. However, the control structures are indeed assessed at this level, to demonstrate the structure selection capability of the proposed MINLP solver. The optimum steady state structure found in this level is

evaluated in the inner-level, which provides the related critical disturbances $\boldsymbol{\theta}^k$ and functional variables $\overline{\mathbf{w}}^k$ for the subsequent outer level.

The second outer-level uses this information to solve its initial NLP relaxation problem, which involves both steady state and DAE models. The corresponding solution initiates the tree enumeration on the binary variables. At the branching point, the binary variables with non-integer solutions are fixed to the binary values according to the branching rules in given section 5.4.3. The subsequent redundancy analysis provides the corresponding functional variables $\overline{\mathbf{w}}^k$ and critical disturbance combinations $\boldsymbol{\theta}^k$ for use in the NLP subproblem. If the subproblem converges to all-integer solutions of the binary variables, the corresponding objective value Φ becomes the lower bound of the optimal solution. Otherwise, the node is fathomed. This enumeration continues until all feasible nodes are accounted for.

5.5.3.3 Results and Discussions

The dynamic MINLP framework converges in two major iterations and in 4.11 CPU minutes. The results are shown in Table 5.9. The tree enumeration at the first outer level is given at Table 5.10. Since this level only considers the nominal disturbance combination, it delivers the optimum nominal steady state net profit, and does not perform redundancy analysis. In addition, there is no specific information for control structure. Therefore, any initial structure becomes the selected control structure. The tentative structure at this stage is $Cool_1$-m_{c1} and $Cool_2$-m_{c2}, and the definite controller structure is determined through dynamic assessment at the second outer-level.

Table 5.9 Superstructure 1: Controllability Assessment Results

	Steady state optimum (iter #1)	Closed loop optimum (iter #2)
Φ \$/h	90.35	87.77
PVAR \$	16.29	15.08
Q_{f1} m^3/s	0.355 (1)	0.324 (1)
Q_{f2} m^3/s	0.206 (1)	0.223 (1)
α_1 (y_{c1})	1 (0)	2 (1)
α_2 (y_{c2})	1 (0)	2 (0)
Variable groups	[T_1, T_2, $Cool_1$, C_2] [$Cool_2$]	[T_2, $Cool_1$, C_2] [T_1, $Cool_2$]
$\overline{\mathbf{w}}^k$	T_1, $Cool_2$	C_2, T_1
θ^k [T_f, C_f]	[315, 21], [298,19.5]	[315, 21], [298,19.5]

Table 5.10 Superstructure 1: Tree Enumeration in the First Outer-Level

Node no		0	1	2
Initial values	y_{c1}	0	0	0
	y_{c2}	0	0	0
Generated results	y_{c1}	0	1	0
	y_{c2}	0	0	1
	Φ \$/h	**90.35**	**90.35**	**90.35**

Optimum solution Fathomed

For the tentative structure, the redundancy analysis in the first inner-level finds the functional variables T_1, $Cool_2$, and the critical disturbance combinations [315,21] and [298,19.5], respectively. These are sent to the second outer-level. The enumeration process and the generated functional variables and critical disturbances are given in Table 5.11. The solutions of NLP relaxation are integers on y_{c1} and y_{c2}, therefore set the lower bound of the dynamic MINLP solution. The second node also finds the integer solutions of y_{c1} and y_{c2}, and this time increases the optimum lower bound. The third node fails to meet the bound, hence fathomed.

Interestingly, the solutions of NLP subproblems in Table 5.11 are similar to the fixed structure cases in Chapter 3. These results relate to the different constraints on $Cool_1$ and $Cool_2$. As explained in Chapter 3, better profitability is achieved by operating T_1, T_2, $Cool_1$ and $Cool_2$ as close as possible to the respective constraints. However, the pairs T_1-$Cool_1$ and T_2-$Cool_2$ are highly interactive. Controlling T_1 or T_2 leave offsets on $Cool_1$ or $Cool_2$ and vice versa. Furthermore, the tighter control tuning and the higher operational values of the controlled variables, the higher the offsets of uncontrolled counterparts, as these become more sensitive to disturbances.

The constraint on $Cool_1$ is relatively relaxed, and that although T_1 is tightly controlled and operated as close as possible to the constraints, all $Cool_1$ profiles remain feasible. This condition does not exist in the second reactor, because controlling T_2 may force $Cool_2$ to violate its tight constraint limit. The feasibility is only guaranteed if T_2 is controlled and operated at lower value, which is not profitable, nor worth controlling. Better profit is achieved in the second reactor only if $Cool_2$ is controlled close to the constraint, since the associated T_2 profiles remain always feasible. This explains the optimality of T_1-m_{c1} and $Cool_2$-m_{c2} structure, as well as why the optimal tuning of T_1-m_{c1} and T_2-m_{c2} structure resulted in the single-PI control structure.

Table 5.11 Superstructure 1: Tree Enumeration in the Second Outer-Level

Node no		0	1	2
Supplied initial values	Variable group	[T_1, T_2, $Cool_1$, C_p], [$Cool_2$]	[T_1, T_2, $Cool_1$, $Cool_2$, C_p],	[T_1, $Cool_1$], [T_2, $Cool_2$], [C_p]
	$\overline{\mathbf{w}}^k$	T_1, $Cool_2$	T_1	T_1, $Cool_2$, C_p
	$\boldsymbol{\theta}^k$ [T_f, C_f]	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [315, 19.5], [298, 21], [298,19.5]
	y_{c1} (α_1)	0 (1)	1 (0)	1 (2)
	y_{c2} (α_2)	0 (1)	0 (0)	0 (2)
Generated results	y_{c1} (α_1)	0 (0)	1 (2)	1 (2)
	y_{c2} (α_2)	0 (0)	0 (2)	1 (0)
	Φ \$/h	46.86	**87.77**	**79.27**

Optimum solution Fathomed

In the $Cool_1$-m_{c1} and $Cool_2$-m_{c2} control strategy, the controlled $Cool_1$ causes large offsets on T_1 profiles, such that both variables are only feasible if operated at the open-loop values. It leads to low values of T_2 and $Cool_2$ as well, because there is not enough energy coming from the first reactor to push the condition in the second reactor any closer to the constraints. The values are so low, that no actual control is required to keep them feasible. Therefore, this control structure performs just as poorly as the open-loop strategy.

The functional variables and critical disturbance combinations in Table 5.11 are generated without any problem, especially regarding the corresponding convex-hull computation. They also comply with the results in Chapter 4, therefore validate the embedded redundancy analysis and elimination within this dynamic MINLP solver.

5.5.4 Process and Controller Structure Selection

This case study involves both process and controller structure selection of the superstructure given in Figure 5.5, which is taken from Bahri (1996) with modification on the controller structure. The dynamic variables are similar to the previous cases, which are the product compositions C_1, C_2 and temperatures T_1, T_2. The feasible operating conditions are defined by the temperatures T_1, T_2, the amount of heat transfer between coolant and reactor $Cool_1$, $Cool_2$ and the final product composition C_p. The manipulated variables, which are coolant flow rates m_{c1} and m_{c2}, may control either T_1 or $Cool_1$ and T_2 or $Cool_2$, respectively.

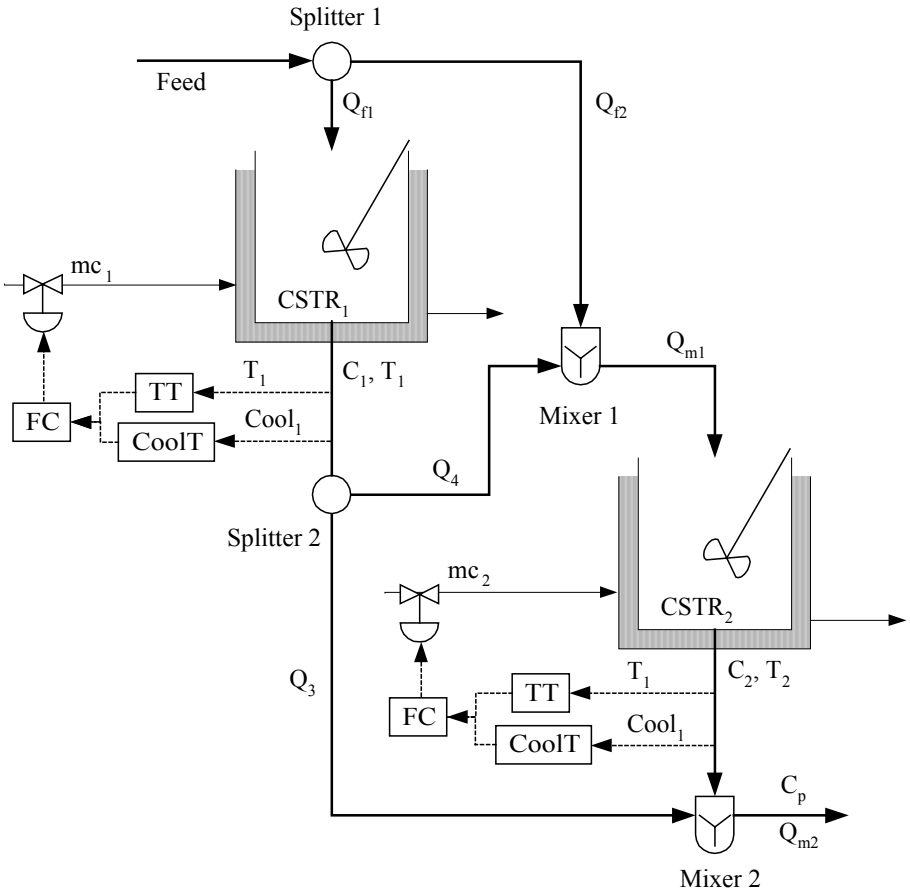


Figure 5.5 Superstructure 2: Process and Controller Structure Selection

The controller structure problem in this case is similar to the previous case study. This time, the new process structure problem is added. The problem is to decide the existence of the flow rates Q_{f1}, Q_{f2}, Q_3 and Q_4. At least one of Q_{f1} or Q_{f2}, and one of Q_3 or Q_4 should exist, which leads to series, parallel, or combination process configuration. These decisions are facilitated by assigning four additional binary variables y_{p1}, y_{p2}, y_{p3}, and y_{p4} to the respective flow rates. Furthermore, the effect of reactor volumes V_1 and V_2 to both process and controller structures is evaluated.

For these purposes, the superstructure is modeled as follows:

$$V_1 \frac{dC_1}{dt} = -k_o e^{-Ec/RT_1} C_1 V_1 + Q_{f1}(C_f - C_1) \tag{5.29}$$

$$V_1 \frac{dT_1}{dt} = D_h k_o e^{-Ec/RT_1} C_1 V_1 + Q_{f1}(T_f - T_1) + Cool_1 \tag{5.30}$$

$$Cool_1 = \frac{U_a m_{c1} cp}{U_a + m_{c1} cp}(T_{c1} - T_1) \tag{5.31}$$

$$Q_3 = Q_{f1} + Q_{f2} \tag{5.32}$$

$$Q_{m1} = Q_3 + Q_{f2} \tag{5.33}$$

$$C_{m1} = \frac{(Q_{f1} C_1 + Q_{f1} C_f)}{Q_{m1}} \tag{5.34}$$

$$T_{m1} = \frac{(Q_{f1} T_1 + Q_{f1} T_f)}{Q_{m1}} \tag{5.35}$$

$$V_2 \frac{dC_2}{dt} = -k_o e^{-Ec/RT_2} C_2 V_2 + Q_{m1}(C_{m1} - C_2) \tag{5.36}$$

$$V_2 \frac{dT_2}{dt} = D_h k_o e^{-Ec/RT_2} C_2 V_2 + Q_{m1}(T_{m1} - T_2) + Cool_2 \tag{5.37}$$

$$Cool_2 = \frac{U_a m_{c2} cp}{U_a + m_{c2} cp}(T_{c2} - T_2) \tag{5.38}$$

$$Q_{m2} = Q_{m1} + Q_4 \tag{5.39}$$

$$C_p = \frac{(Q_4 C_1 + Q_{m1} C_2)}{Q_{m2}} \tag{5.40}$$

$$\Phi_N = 10(Q_{f1}C_{fN} + Q_{f2}C_{fN} - 0.3(Q_{f1} + Q_{f2})) - 0.01Cool_1 \ldots \\ -0.1Cool_2 - 0.1(Q_{f1} + Q_{f2}) \tag{5.41}$$

$$e_{11} = y_{c1}(T_{1ss} - T_1) \tag{5.42}$$

$$\frac{dI_{11}}{dt} = e_{11} \tag{5.43}$$

$$e_{12} = (1 - y_{c1})(Cool_{1ss} - Cool_1) \tag{5.44}$$

$$\frac{dI_{12}}{dt} = e_{12} \tag{5.45}$$

$$m_{c2} = m_{c2ss} + \alpha_1\left(K_{c11}\left(e_{11} + \frac{I_{11}}{\tau_{11}}\right) + K_{c22}\left(e_{12} + \frac{I_{12}}{\tau_{12}}\right)\right) \tag{5.46}$$

$$e_{21} = y_{c2}(T_{2ss} - T_2) \tag{5.47}$$

$$\frac{dI_{21}}{dt} = e_{21} \tag{5.48}$$

$$e_{22} = (1 - y_{c2})(Cool_{2ss} - Cool_2) \tag{5.49}$$

$$\frac{dI_{22}}{dt} = e_{22} \tag{5.50}$$

$$m_{c2} = m_{c2ss} + \alpha_2\begin{pmatrix} K_{c21}\left(e_{21} + \frac{I_{21}}{\tau_{21}}\right) + K_{c22}\left(e_{22} + \frac{I_{22}}{\tau_{22}}\right) + \ldots \\ K_{int\,1}\left(e_{11} + \frac{I_{11}}{\tau_{int\,1}}\right) \end{pmatrix} \tag{5.51}$$

This superstructure is assessed for the open-loop and closed-loop conditions. The closed-loop uses a pre-designed multivariable controller using the DICOF package (Agamennoni and Romagnoli, 1993), which is realisable on a well-known Proportional-Integral (PI) controller structure. The optimum controller parameters are tuned further with the scaling factors α_1 and α_2. The binary assignments related to controller and process structure are given in Tables 5.12 and 5.13 respectively.

Table 5.12. Controller Parameters and Logical Assignments

Control loop		Binary	Proportional	Reset
i	Control pair	Assignment	gain K_i	times τ_I
11	T_1 - m_{c1}	$y_{c1} = 1$	-3.876	9.346
12	$Cool_1$ - m_{c1}.	$y_{c1} = 0$	3.79	6.135
21	T_2 - m_{c2}	$y_{c2} = 1$	-3.876	9.346
22	$Cool_2$ - m_{c2}.	$y_{c2} = 0$	3.79	6.135
int1	T_1 - m_{c2}	$y_{c2} = 0$	0.0013	1.5

idx = controller parameter index

Table 5.13. Process Structure and Binary Assignments

Binary assignments	Decision	Binary assignments	Decision
$y_{p1} = 0$	Q_{f1} off	$y_{p1} = 1$	Q_{f1} on
$y_{p2} = 0$	Q_{f2} off	$y_{p2} = 1$	Q_{f2} on
$y_{p3} = 0$	Q_3 off	$y_{p3} = 1$	Q_3 on
$y_{p4} = 0$	Q_4 off	$y_{p4} = 1$	Q_4 on

The superstructure is optimised to achieve the optimum nominal net profit similar to the previous case. Reflecting the structural problems, the design variables consist of Q_{f1}, Q_{f2}, Q_3, Q_4, α_1, α_2, V_1, V_2. y_{c1}, y_{c2}, y_{p1}, y_{p2}, y_{p3} and y_{p4} as well as the initial conditions of T_{1ss}, T_{2ss}, C_{1ss}, C_{2ss}, $Cool_{1ss}$, and $Cool_{2ss}$.

The associated constraints for these design variables are given in (5.52).

$$
\begin{aligned}
&g_1 : T_1 \le 350 && g_2 : T_2 \le 350 \\
&g_3 : Cool_1 \le 30 && g_4 : Cool_2 \le 20 \\
&g_5 : C_2 \le 0.3 && g_6 : Q_{f1} + Q_{f2} \le 0.8 \\
&g_7 : Q_{f1} \ge 0.05 y_{p1} && g_8 : Q_{f1} \le 0.8 y_{p1} \\
&g_9 : Q_{f2} \ge 0.05 y_{p2} && g_{10} : Q_{f2} \le 0.8 y_{p2} \\
&g_{11} : Q_3 \ge 0.05 y_{p3} && g_{12} : Q_3 \le 0.8 y_{p3} \\
&g_{13} : Q_4 \ge 0.05 y_{p4} && g_{14} : Q_4 \le 0.8 y_{p4} \\
&g_{15} : 0 \le \alpha_1 && g_{16} : \alpha_1 \le 2 \\
&g_{17} : 0 \le \alpha_1 && g_{18} : \alpha_1 \le 2 \\
&g_{19} : V_1 \ge 4.5 && g_{20} : V_1 \le 6 \\
&g_{21} : V_2 \ge 4.5 && g_{22} : V_2 \le 6 \\
&g_{23} : y_{p1} + y_{p2} \ge 1 && g_{24} : y_{p3} + y_{p4} \ge 1
\end{aligned}
\tag{5.52}
$$

5.5.4.1 Framework Implementation

Similar to the previous cases, the models, objective, and constraints, are arranged according to the formulation (5.8)-(5.9). The open-loop superstructure model consists of (5.29)-(5.41), while the closed-loop model consists of (5.29)-(5.51). The initial outer-level solves the steady state version of the respective model, the inner-levels solve the full DAE model, and the second outer-levels solve both the steady state and DAE models. All DAEs are simulated over 100 seconds time horizon with 1-second intervals, and the outputs are $Cool_1$, $Cool_2$, m_{c1}, m_{c2}, and Φ_N; in addition to the state variables T_1, T_2, C_1, C_2, I_{11}, I_{12}, I_{21}, and I_{22}.

The constraints (5.52) apply to both open-loop and closed-loop cases. Constraints $\mathbf{g_7}$-$\mathbf{g_{14}}$ relate the binary decisions to the flow-rate existence. These relationships guarantee the amounts of flow-rate if y_p=1, and none otherwise. At least one of Q_{f1} or Q_{f2} and one of Q_3 or Q_4 should exist, and these are represented by $\mathbf{g_{23}}$-$\mathbf{g_{24}}$, which are evaluated in the integer feasibility test.

Similar to the case study in section 5.5.3, the integer feasibility test in the first outer level provides the tentative controller structure for the subsequent levels. The same test within the second outer level is followed by the redundancy analysis and elimination procedure to provide the functional variables $\overline{\mathbf{w}}^{\mathbf{k}}$ and critical disturbance combination $\boldsymbol{\theta}^{k}$ for the corresponding NLP subproblem.

The combination of the constraints $\mathbf{g_{23}}$-$\mathbf{g_{24}}$ (5.52) and the binary variables ($\mathbf{y_{p1-4}}$ for the open-loop case, $\mathbf{y_{p1-4}}$ and $\mathbf{y_{c1-2}}$ for the closed-loop case) give rise to a 36 nodes enumeration tree at the dynamic MINLP problem. Each node of the tree involves one combination of process and controller structure, and the associated functional variables and critical disturbance combinations.

5.5.4.2 Results and Discussions

Both open-loop and closed-loop cases converge in two major iterations. The controllability assessment and the tree enumeration results are shown in Table 5.14 to 5.19. The optimum process and controller structures are shown in Figures 5.6 and 5.7, and the respective dynamics due to $\boldsymbol{\theta}^{k}$ are given in Figures 5.8 and 5.9.

The open-loop controllability assessment completes in two iterations and 1.636 CPU minutes. The results are shown in Table 5.14. The first outer-level of the open-loop case is the steady state MINLP problem. The associated tree enumeration is shown in Table 5.15. Here, the solution of the NLP relaxation problem (node 0) provides the highest profit, but not all of involved binary variables $\mathbf{y_p}$ have integer values. Therefore, the lower bound of the profit is determined by the solution of NLP subproblem at the first node, where the profit value is maintained, and all $\mathbf{y_p}$ are integer. The other nodes do not give better profits or all integer $\mathbf{y_p}$. Therefore, they are fathomed. The optimum process structure found in this level is the parallel configuration, and this structure is maintained, regardless of the priorities assigned to $\mathbf{y_p}$. This consistency is specifically due to the connection between y_{p1} and y_{p2} with the respective flow-rates Q_{f1}, Q_{f2}, which directly determines the profit.

Table 5.14 Open-Loop Superstructure 2: Controllability Assessment Results

	Steady state optimum (iter #1)	Open-loop optimum (iter #2)
Φ \$/h	109.410	77.2919
PVAR \$	318.091	271.242
Q_{f1} m^3/s (y_{p1})	0.356 (1)	0.252 (1)
Q_{f2} m^3/s (y_{p2})	0.302 (1)	0.231 (1)
Q_3 m^3/s (y_{p3})	0.000 (0)	0.000 (0)
Q_4 m^3/s (y_{p4})	0.356 (1)	0.252 (1)
V_1 m^3	4.500	5.864
V_2 m^3	4.500	5.356
Variable groups	[T_1, T_2, $Cool_1$, C_2, $Cool_2$]	[T_2, $Cool_1$, C_2], [T_1, $Cool_2$]
$\overline{\mathbf{w}}^k$	T_1	C_2, T_1
θ^k [T_f, C_f]	[315, 21], [298,19.5]	[315, 21], [298,19.5]

Table 5.15 Open-Loop Superstructure 2: Tree Enumeration in the First Outer-Level

	Node	0	1	2	3	4
Initial values	y_{p1}	0	0.515	0	1	1
	y_{p2}	0	0.515	1	0	1
	y_{p3}	0	0	0	0	1
	y_{p4}	0	1	1	1	1
Generated results	y_{p1}	0.515	1	0.44	1	1
	y_{p2}	0.515	1	1	0	1
	y_{p3}	0	0	0	0	1
	y_{p4}	1	1	1	1	1
	Φ \$/h	109.41	**109.41**	**109.41**	**69.58**	**106.81**

Optimum solution Fathomed

The second outer-level, as shown in Table 5.16, retains the parallel structure, but with higher reactor volumes V_1 and V_2. These increases reduce the process sensitivity to the disturbances, specifically the upsets on T_1, T_2, C_1 and C_2, and therefore increase the respective set points. However, the upsets of $Cool_2$ limit the increase of T_2 set point and V_2. This determines the corresponding values of T_1, $Cool_1$, and V_1. Despite the increased volumes, the net profit still drops by 29.36% from the steady state optimum. The cumulative profit variation over the time horizon is 9.71% of the hourly net profit, which is a serious reason for controller implementation.

Table 5.16 Open-Loop Superstructure 2: Tree Enumeration in the Second Outer-Level

Node no		0	1	2
Supplied initial values	Variable group	[T_1, T_2, $Cool_1$, $Cool_2$, C_p]	[T_1, T_2, $Cool_1$, $Cool_2$, C_p]	[T_1, T_2, $Cool_1$, $Cool_2$, C_p]
	$\overline{\mathbf{w}}^k$	T_1	C_p	C_p
	$\boldsymbol{\theta}^k$	[315, 21] [298, 19.5]	[315, 21] [298, 19.5]	[315, 21] [298, 19.5]
	y_{p1}	1	0	1
	y_{p2}	1	1	0
	y_{p3}	0	0	0
	y_{p4}	1	1	1
Generated results	y_{p1}	1	0	1
	y_{p2}	1	1	0
	y_{p3}	0	-0.0047	-0.0013
	y_{p4}	1	0.9997	0.9994
	Φ \$/h	**77.2919**	**73.87.47**	**76.9657**

Optimum solution Fathomed

The closed-loop case involves the binary variables $\mathbf{y_p}$ and $\mathbf{y_c}$ for process and controller structure selection respectively. The process structure selected at the first outer-level is the parallel configuration identical to the open-loop case. Since there is no specific information to determine the control structure at this level, any initial controller structure becomes the tentative control structure, similar to the case study in section 5.5.3. Tables 5.17-5.18 show the tentative control structure selected for the given initial $\mathbf{y_c}$ are $Cool_1$-m_{c1} and $Cool_2$-m_{c2}. The definite optimum control structure is determined through dynamic MINLP in the second outer-level.

Table 5.17 Closed-Loop Superstructure 2: Controllability Assessment Results

	Steady state optimum (iter #1)	Open-loop optimum (iter #2)
Φ \$/h	109.410	106.536
PVAR \$	318.091	9.607
Q_{f1} m^3/s (y_{p1})	0.356 (1)	0.356 (1)
Q_{f2} m^3/s (y_{p2})	0.302 (1)	0.288 (1)
Q_3 m^3/s (y_{p3})	0.000 (0)	0.052 (1)
Q_4 m^3/s (y_{p4})	0.356 (1)	0 304 (1)
α_1 (y_{c1})	1 (0)	2 (1)
α_2 (y_{c2})	1 (0)	2 (0)
V_1 m^3	4.500	4.500
V_2 m^3	4.500	4.500
Variable groups	[T_1, T_2, $Cool_1$, C_2] [$Cool_2$]	[T_2, $Cool_1$, C_2] [T_1, $Cool_2$]
$\overline{\mathbf{w}}^k$	T_1, $Cool_2$	C_2, T_1
θ^k [T_f, C_f]	[315, 21], [298,19.5]	[315, 21], [298,19.5]

Table 5.18 Closed-loop Superstructure 2: Tree Enumeration in the First Outer-Level

	Node	0	1	2	3
Supplied initial values	y_{p1}	0	1	0	1
	y_{p2}	0	1	1	0
	y_{p3}	0	0	0	0
	y_{p4}	0	1	1	1
	y_{c1}	0	0	0	0
	y_{c2}	0	0	0	0
Generated results	y_{p1}	0.515	1	0	1
	y_{p2}	0.515	1	1	0
	y_{p3}	0	0	0.884	0
	y_{p4}	1	1	0.097	1
	y_{c1}	0	0	0	0
	y_{c2}	0	0	0	0
	Φ $/h	109.41	**109.41**	**47.34**	**69.58**

Optimum solution Fathomed

The first inner-level provides the required $\overline{\mathbf{w}}^{\mathbf{k}}$ and $\boldsymbol{\theta}^{k}$ for the initial NLP relaxation of the second outer-level. The tree enumeration results of the latter level are shown in Table 5.19, which finally delivers T_1-m_{c1} and $Cool_2$-m_{c2} controller structure. The respective net profit improves up to 37.85% compared to the open loop case, and maintains 0.007% profit variations, which is a significant incentive of process control implementation. The controllers tightly reject the disturbances and do not require more help from extra volumes. Therefore, the reactor volumes are kept at the minimum.

Table 5.19 Closed-Loop Superstructure 2: Tree Enumeration in the Second Outer-Level

Node no		0	1	2	3	4	5	6
Supplied initial values	$\overline{\mathbf{w}}^k$	$Coo1_1$, $Cool_2$	T_1, $Cool_2$	T_1	T_1, $Cool_2$	T_2	T_1	T_2, $Cool_2$, C_p
	$\boldsymbol{\theta}^k$	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [298,19.5]	[315, 21], [298,19.5]
	y_{p1}	1	1	1	1	1	1	1
	y_{p2}	1	1	1	1	1	1	1
	y_{p3}	0	0	0	1	1	1	0
	y_{p4}	1	1	1	1	0	0	1
	y_{c1}	0	1	0	1	1	0	1
	y_{c2}	0	0	0	0	0	0	1

Optimum solution

Fathomed

Table 5.19 Closed-loop superstructure 2: tree enumeration in the second outer-level (continued)

Node no		0	1	2	3	4	5	6
Generated results	y_{p1}	1	1	1	1	1	1	1
	y_{p2}	1	1	1	1	1	1	1
	y_{p3}	0	0.0768	0	1	1	1	0
	y_{p4}	1	1	1	1	0	0	1
	y_{c1}	1	1	1	1	1	0	1
	y_{c2}	0.0292	0	0	0	0	0	1
	Φ $/h	104.7834	106.5358	**84.8888**	**106.5358**	**90.4424**	**86.1548**	**92.30110**

Optimum solution

Fathomed

The parallel configuration is also dominant in the final closed-loop structure. However, a small amount of the first reactor product is passed to the second one (Q_3=0.052m^3/s). Note that the temperature of Q_3 is T_1, which is higher than the maximum feed temperature T_f. Therefore, although small, T_2 and $Cool_2$ are increased, as well as Q_{f2} conversion. This leads to a slightly higher profit in comparison to purely closed-loop parallel structure.

It should be noted that this superstructure involves reactors with different specifications. If all reactors have identical specifications, theoretically the serial configuration is preferable. This configuration improves the conversion and consumes less energy. However, if any reactor in the superstructure fails, the problem propagates along the superstructure and severely degrades the overall process. On the other hand, parallel configuration is more robust. It minimises the interactions, prevents disturbance propagation between reactors, and allows easier control of each reactor. However, the conversion rate may be lower than the serial configuration. Therefore, the optimum process controllability and profitability in consideration of individual reactor specification may involve some trade-off between parallel and serial structures.

Overall, this case study demonstrates the interesting interaction between process and controller structure. The steady state optimum structure apparently does not necessarily produce the dynamic optimum structure. Furthermore, the optimum structure for open loop dynamics is not necessarily the same as the optimum structure for the closed loop dynamics. It shows that the open loop analysis does not provide complete information to determine the optimum control structure, and emphasises the importance of considering the process dynamics in process design and synthesis. In summary, the process controllability depends on both process and controller characteristics.

Following the process syntheses of academic case studies, it would be interesting to evaluate the framework on a real industrial case. This particular application, with higher complexities at DAE and structure formulations is given in Chapter 7.

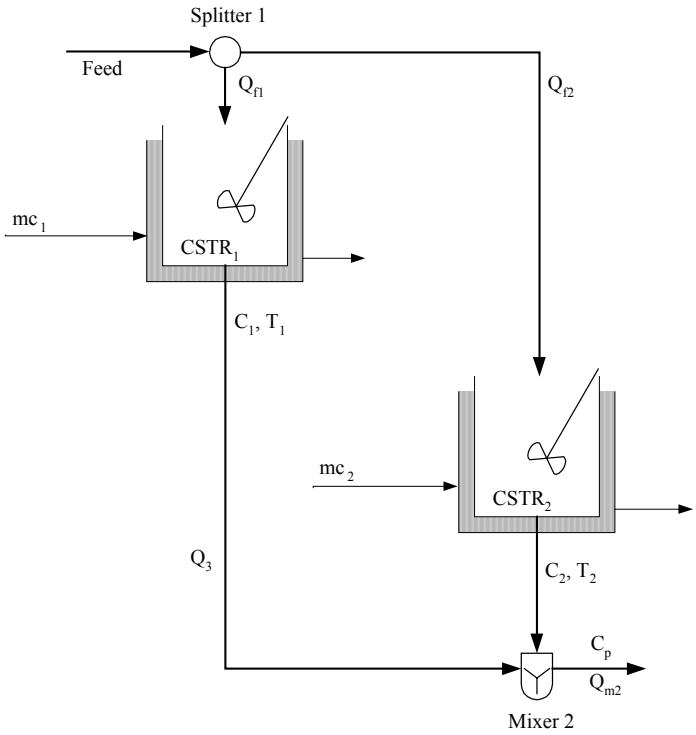


Figure 5.6 Open-Loop Superstructure 2: Optimum Process Structure

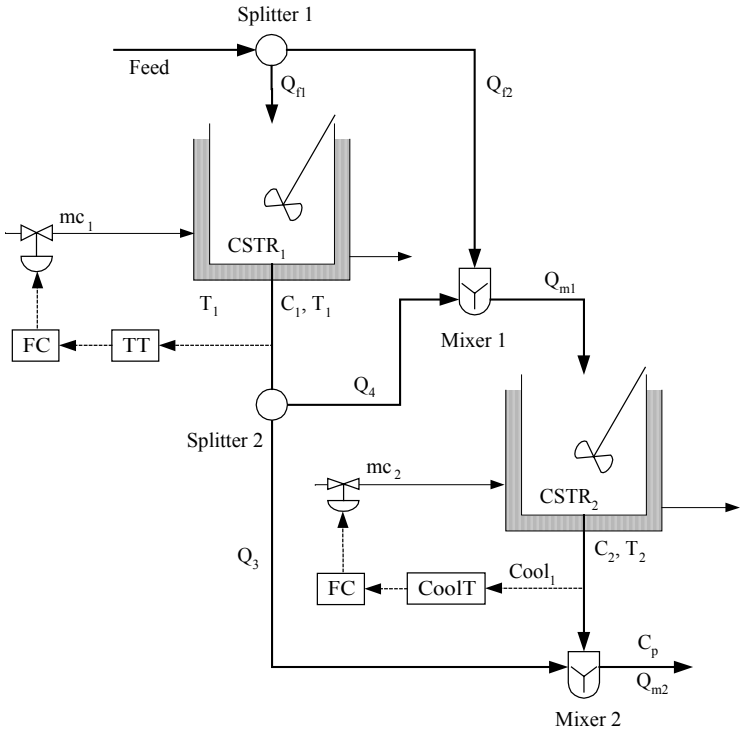


Figure 5.7 Closed-Loop Superstructure 2: Optimum Process and Controller Structure

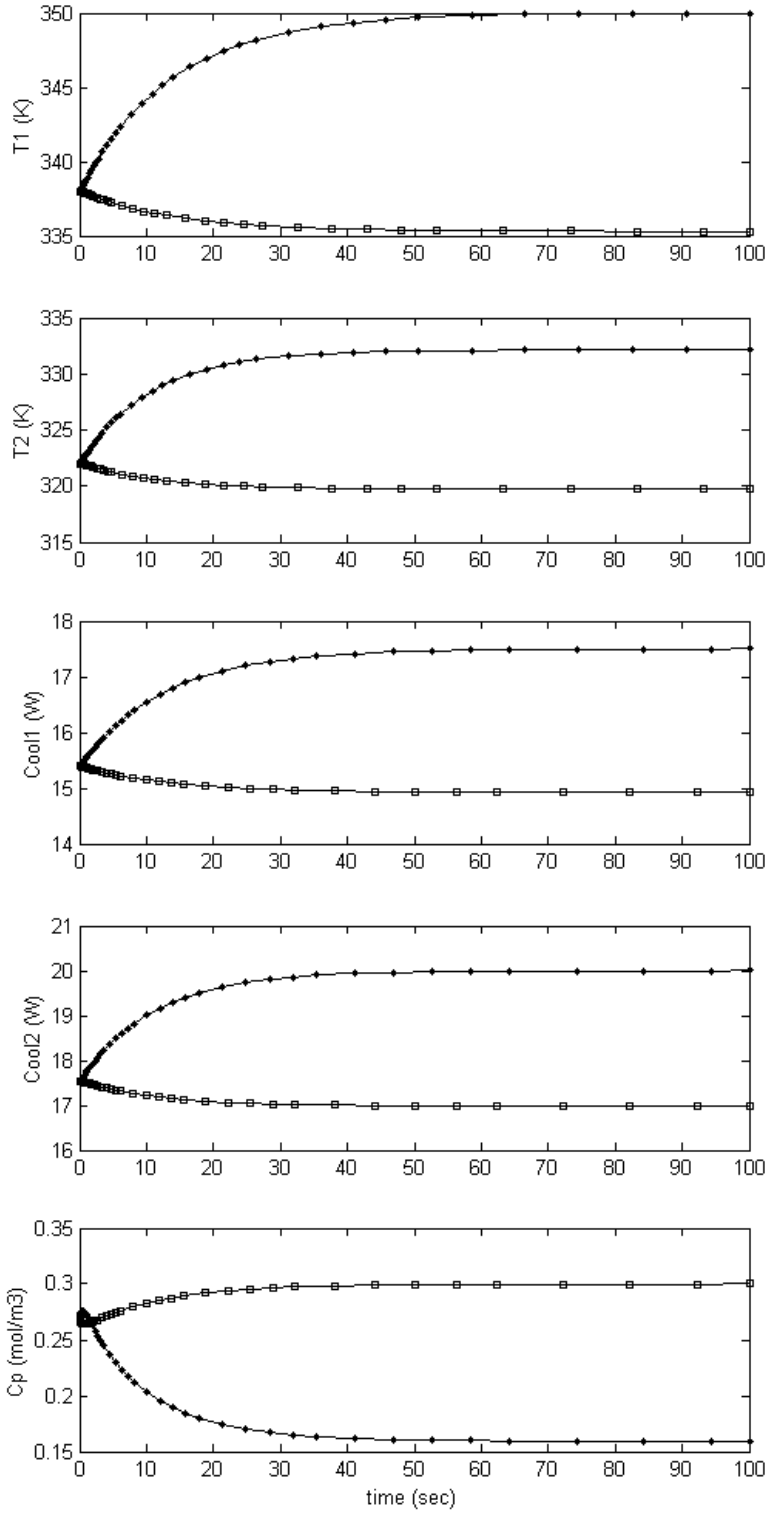


Figure 5.8 Open-Loop Superstructure 2: Dynamic Responses of T_1, T_2, $Cool_1$, $Cool_2$ and C_p due to θ^k

-*- : T_f = 315K, C_f = 21mol/m^3, -□- : T_f = 298K, C_f = 19.5mol/m^3

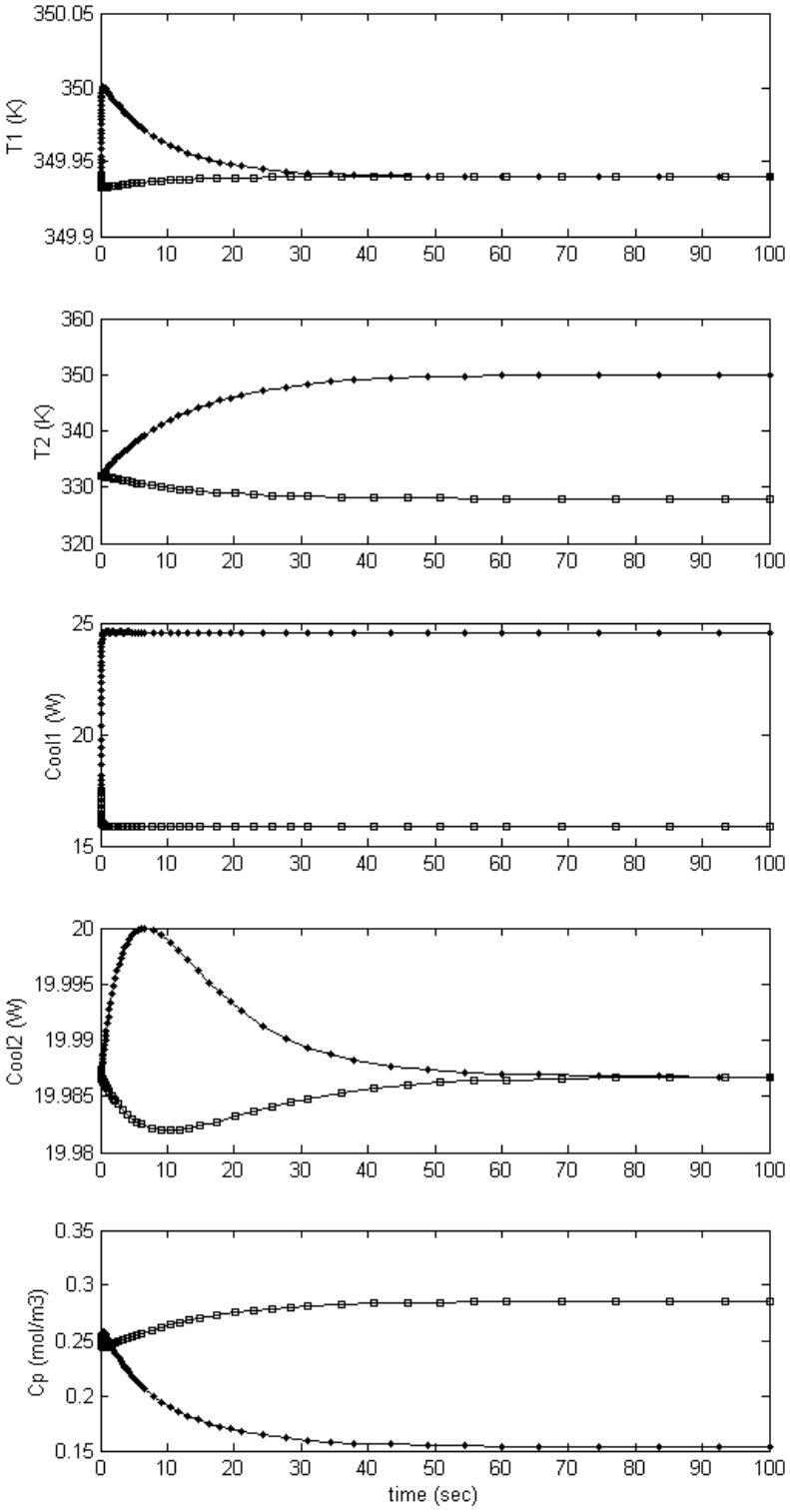


Figure 5.9 Closed-Loop Superstructure 2: Dynamic Responses of T_1, T_2, $Cool_1$, $Cool_2$ and C_p due to $\boldsymbol{\theta}^k$

-*- : T_f = 315K, C_f = 21mol/m^3, -□- : T_f = 298K, C_f = 19.5mol/m^3

5.6 Conclusion

The process synthesis problem within the proposed Dynamic Operability Framework has been addressed. For this purpose, the framework maintains its two-level approach, where the outer level this time becomes the dynamic semi-infinite MINLP problem.

The dynamic MINLP solution has been developed based on the Branch and Bound method. The features include the accommodation of redundancy analysis and elimination procedure described in Chapter 4, the compact MINLP superstructure formulation, the combined depth-search and breadth-search, the accompanying branching rules, and the convergence treatment of the NLP solutions.

The performance of the algorithm has been demonstrated in four case studies. Fast and efficient convergence has been shown for convex and steady state problems, while the global optimum solutions have been found so far on the dynamic problems. The tractability of the problem is theoretically limited by the number of binary variables, the number of functional variables and the DAE index. Further and detailed analysis to complement this development is still required, and therefore recommended for future work.

The application on the dynamic and uncertain process superstructure show strong interaction between process and controller characteristics. This emphasises that process controllability depends on both the process and the controller.

The application of this proposed framework to real industrial case is demonstrated in Chapter 7.

5.7 Nomenclature

Acronyms		
AOS_θ	:	Achievable Output Space due to disturbances
BB	:	Branch and Bound
CNF	:	Conjunctive Normal Form
CRP	:	Convex-hull Relaxation Problem
DAE	:	Differential Algebraic Equations
DOF	:	Dynamic Operability Framework
DOS	:	Desired Output Space
EDS	:	Expected Disturbance Space
GBD	:	Generalised Benders Decomposition
GDP	:	Generalised Disjunctive Programming
HEN	:	Heat Exchanger Network
ICI	:	Integral Controllability with Integrity
IS	:	Integral Stabilisability
QP	:	Quadratic Programming
OA	:	Outer Approximation
OA/ER	:	Outer Approximation with Equality Relaxation
OA/ER/AP	:	Outer Approximation with Equality Relaxation and Augmented Penalty
r-OCI	:	Regulatory Output Controllability Index
MIDO	:	Mixed Integer Dynamic Programming
MILP	:	Mixed Integer Linear Programming
MINLP	:	Mixed Integer Nonlinear Programming
MIQP	:	Mixed Integer Quadratic Programming
NLP	:	Nonlinear Programming
PI	:	Proportional – Integral
PVAR	:	Profit Variation
SIDMINLP	:	Semi-Infinite Dynamic Mixed Integer Nonlinear Programming
SQP	:	Sequential Quadratic Programming

Variables and Matrices		
C	:	Concentration of A in CSTR (mol/m^3)
Cool	:	Amount of heat removed from CSTR by coolant
c_p	:	Heat capacity (J/Kg K)
Dh	:	Heat of reaction (°Km^3/mol)
E	:	Set of equations
Ec	:	Activation energy
e	:	Error variables
$\mathbf{g_j}$	:	Inequality constraints
$\mathbf{g_{jd}}$	:	Logic propositions
g_o	:	Function of process parameters
$\mathbf{h_i}$	:	Equality constraints
I	:	Integral of error variables
Ir	:	Set of reduced inequality
IL	:	Set of previously found integer solutions
i	:	Index of equality constraint
i_{DAE}	:	Index of a set of Differential Algebraic Equations
j	:	Index of inequality constraint
jd	:	Index of logic propositions
K	:	Proportional gain of PI controller
k	:	Binary counter
k_o	:	Reaction constant
L	:	Set of lower bound values
M	:	Large scalar values to activate or deactivate constraints
m_c	:	Coolant flow rate (m^3/sec)
n_{yfeas}	:	Number of feasible binary combination
n_{iter}	:	Number of iterations of an NLP
$n_{maxiter}$	:	Number of maximum iterations of an NLP
n_w	:	Number of functional variables
n_{xw}	:	Total number of dynamic variables
n_y	:	The number of binary variables
O	:	Order of computational complexity
p	:	Vector of process parameters
Q	:	Liquid Flow rate (m^3/sec)
R	:	Reaction rate (moles/sec)
T	:	Temperature of the reactor (°K)
t	:	Time (sec)
t_{DAE}	:	Time required to solve one DAE (sec)
U	:	Set of upper bound values
u	:	Vector of manipulated variables
Ua	:	Overall heat transfer coefficient (W/°K)
V	:	CSTR volume (m^3)

Variables and Matrices		
$\mathbf{W}$	:	Set of measurement variable values
$\mathbf{w}$	:	Vector of measured variables
$\overline{\mathbf{w}}$	:	Vector of functional measurement variables
$\mathbf{X}$	:	Set of state variable values
$\mathbf{x}$	:	Vector of state variables
$\mathbf{Y}$	:	Set of binary/logic variables
$\mathbf{y}$	:	Vector of binary/logic decision variables
$\mathbf{Z}$	:	Set of continuous decision variable values
$\mathbf{z}$	:	Vector of continuous decision variables
$\overline{\mathbf{z}}$	:	Vector of augmented decision variables

Greek letters		
α	:	Scaling factors for controller parameters
$\mathbf{\Gamma}$	:	Set of all disturbance realisations
Φ	:	Objective function
μ	:	Volume of a multi dimensional space
$\boldsymbol{\theta}$	:	Vector of disturbances and uncertainties
τ	:	Integral/reset time of PI controller
Ω	:	Disjunctive expressions

Superscripts		
k	:	Critical values
l	:	Lower bounds
u	:	Upper bounds
N	:	Nominal values
*	:	Optimal solution from the last iteration

Subscripts		
ss	:	Steady state conditions
1	:	First CSTR
2	:	Second CSTR
f	:	Feed
c	:	Control
m	:	Mixer output
p	:	Process

5.8 References

Adjiman, C. S., Dallwig, S., Floudas, C. A. and Neumaier, A. (1998a). "A Global Optimization Method, α-BB, for General Twice-Differentiable Constrained NLPs - I. Theoretical Advances", Computers and Chemical Engineering, 22 (9), 1137-1158.

Adjiman, C. S., Dallwig, S., Floudas, C. A. and Neumaier, A. (1998b). "A Global Optimization Method, α-BB, for General Twice-Differentiable Constrained NLPs - II. Implementation and Computational Results", Computers and Chemical Engineering, 22 (9), 1159-1179.

Agamennoni, O. and Romagnoli, J. A. (1993). "DICOF: A Package for Multi-Variable Controller Design in Frequency Domain", CACHE News, 36, 15.

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2000). "Flexibility Analysis and Design of Linear Systems by Parametric Programming", AIChE Journal, 46 (2), 335-354.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2002a). "A Case Study in Simultaneous Design and Control Using Rigorous, Mixed-Integer Dynamic Optimization Models", Industrial and Engineering Chemistry Research, 41, 760-778.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2002b). "Flexibility Analysis and Design Using a Parametric Programming Framework", AIChE Journal, 48 (12), 2851-2686.

Bansal, V., Ross, R., Perkins, J. D. and Pistikopoulos, E. N. (1998). "Interactions of Design and Control: Double Effect Distillation Systems", IFAC, Corfu, Greece.

Barber, C. B. (1996). "Qhull", Minneapolis, USA, The Geometry Center, University of Minnesota.

Bemporad, A., Ferrari-Trecate, G. and Morari, M. (1999). "Observability and Controllability of Piecewise Affine and Hybrid Systems", Institut fur Automatik, ETH - Swiss Federal Institute of Technology, Zurich, Switzerland.

Borchers, B. and Mitchell, J. E. (1996). "A Computational Comparison of Branch and Bound and Outer Approximation Algorithms for 0-1 Mixed Integer Nonlinear Programs".

Brooke, A. D., Kendrick and Meeraus, A. (1992). "GAMS 2.25 User's Guide".

Campo and Morari, M. (1994). "Achievable Closed-Loop Properties of Systems under Decentralised Control: Conditions Involving Steady-State Gain", IEEE Transactions of on Automatic Control, 39 (5), 932 - 943.

Cervantes, A. and Biegler, L. T. (1997). "Large-Scale DAE Optimization Using Simultaneous Nonlinear Programming Formulations", AIChE Journal, 44, 1038 - 1050.

Dimitriadis, V. D. and Pistikopoulos, E. N. (1995). "Flexibility Analysis of Dynamic Systems", Industrial and Engineering Chemistry Research, 34, 4451 - 4462.

Duran, M. A. and Grossmann, I. E. (1986). "An Outer Approximation Algorithm for a Class of Mixed Integer Nonlinear Programs", Mathematical Programming, 36, 307.

Ekawati, E. (2002). "Systematic Controllability Assessment in Dynamic Operability Framework", The Third Inter University Postgraduate Electrical Engineering Symposium, Murdoch University, Rockingham Campus, Australia.

Fisher, W. R., Doherty, M. F. and Douglas, J. M. (1988). "The Interface between Design and Control. 1. Process Controllability", Industrial and Engineering Chemistry Research, 27, 597-605.

Fletcher, R. and Leyffer, S. (1996). "Solving Mixed Integer Nonlinear Programs by Outer Approximation", Mathematical Programming, 66, 327 - 349.

Floudas, C. A. (1995). "Nonlinear and Mixed Integer Optimization: Fundamentals and Applications", Oxford University Press.

Floudas, C. A., Aggarwal, A. and Ciric, A. G. (1988). "Global Optimum Search for Nonconvex NLP and MINLP Problems", Annual AIChE Meeting, Washington D.C., USA.

Geoffrion, A. m. (1972). "Generalized Benders Decomposition", Journal of Optimization Theory and Application, 10 (4), 237-260.

Grossman, I. E. (2002). "Review of Nonlinear Mixed-Integer and Disjunctive Programming Techniques for Process Systems Engineering", Submitted for publication.

Grossmann, I. E. (1990). "MINLP Optimisation Strategies and Algorithms for Process Synthesis", Dept. Chemical Engineering, Carnegie Mellon University, Pittsburgh.

Grossmann, I. E. and Biegler, L. T. (2003). "Restropective on Optimization", Submitted for publication.

Grossmann, I. E. and Kravanja, Z. (1997). "Large Scale Optimisation with Applications, Part II: Optimal Design and Control", Mixed-Integer Nonlinear Programming: A Survey of Algorithms and Applications, Biegler, L. T., Springer - Verlag, 73 - 100.

Heath, J. A., Kookos, I. K. and Perkins, J. D. (2000). "Process Control Structure Selection Based on Economics", AIChE Journal, 46 (10), 1998 - 2016.

Holmstrom, K., Goran, A. O. and Edvall, M. M. (2003). "User's Guide for Tomlab/MINLP", Tomlab Optimization A B.

Kocis, G. R. and Grossmann, I. E. (1987). "Relaxation Strategy for the Structural Optimization of Process Flowsheets", Industrial and Engineering Chemistry Research, 26, 1869 - 1880.

Lee, S. and Grossman, I. E. (2000). "New Algorithms for Nonlinear Generalized Disjunctive Programming", Computers and Chemical Engineering, 24, 2125-2141.

Leyffer, S. (2001). "Integrating SQP and Branch-and-Bound for Mixed Integer Nonlinear Programming", Computational Optimization and Applications, 18, 295-309.

Mathworks (2000). "Optimisation Toolbox User's Guide Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Mohideen, J., Perkins, J. D. and Pistikopoulos, E. N. (1996). "Optimal Design of Dynamic Systems under Uncertainty", AIChE Journal, 42, 2251-2272.

Narraway, L. and Perkins, J. (1994). "Selection of Process Control Structure Based on Economics", Computers and Chemical Engineering, 18 (Supplementary), S511-S515.

Narraway, L. T., Perkins, J. D. and Barton, G. W. (1991). "Interaction between Process Design and Process Control: Economic Analysis of Process Dynamics", Journal of Process Control, 1, 243 - 250.

Nemhauser, G., Kan, A. R. and Todd, M. (1994). "Optimization", North Holland.

Nishida, N., Stephanopoulos, G. and Westerberg, A. W. (1981). "A Review of Process Synthesis", AIChE Journal, 27, 321 - 351.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994a). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Dynamic and Control Structure Consideration", Industrial and Engineering Chemistry Research, 33, 1738 - 1755.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994b). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Flexibility and Structural Controllability Aspects", Industrial and Engineering Chemistry Research, 33, 1718 - 1737.

Pistikopoulos, E. N. and Grossmann, I. E. (1988a). "Evaluation and Redesign for Improving Flexibility in Linear System with Infeasible Nominal Conditions", Computers and Chemical Engineering, 12, 841.

Pistikopoulos, E. N. and Grossmann, I. E. (1988b). "Optimal Retrofit Design for Improving Process Flexibility in Linear System", Computers and Chemical Engineering, 12, 719 - 731.

Raman, R. and Grossmann, I. E. (1994). "Modeling and Computational Techniques for Logic Based Integer Programming", Computers and Chemical Engineering, 18 (7), 563-578.

Saboo, A. K., Morari, M. and Woodcock, D. C. (1985). "Design of Resilient Processing Plants-VIII: A Resilience Index for Heat Exchanger Networks", Chemical Engineering Science, 40, 1553 - 1565.

Sahinidis, N. (1996). "Baron: A General Purpose Global Optimisation Software Package", Journal of Global Optimisation, 8 (2), 201-205.

Sahinidis, N. V. (2003). "Optimization under Uncertainty: State-of-the-Art and Opportunities", Foundations of Computer-Aided Process Operations, Coral Springs, Florida, USA, CACHE.

Schweiger, C. A. and Floudas, C. A. (1997). "Interaction of Design and Control: Optimization with Dynamic Models", Optimal Control: Theory, Algorithms and Applications, Hager, W. W. and Pardalos, P. M., Kluwer Academic Publishers B. V.

Schweiger, C. A. and Floudas, C. A. (1998). "Process Synthesis, Design and Control: A Mixed-Integer Optimal Control Framework", IFAC Conference on Dynamics and Control of Process Systems, Corfu, Greece, Elsevier.

Solberg, I. (1999). "Fminconset", Mathworks.

Till, J., Engell, S., Panek, S. and Stursberg, O. (2003). "Applied Hybrid System Optimisation: An Empirical Investigation of Complexity", IFAC Conference of Analysis and Design of Hybrid Systems, St. Malo, France.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Vecchietti, A., Lee, S. and Grossmann, I. E. (2003). "Modelling of Discrete/Continuous Optimization Problems: Characterization and Formulation of Disjunction and Their Relaxations", Computers and Chemical Engineering, 27, 433-448.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10 (2), 185 - 194.

Viswanathan, J. and Grossmann, I. E. (1990). "A Combined Penalty Function and Outer Approximation Method for MINLP Optimisation", Computers and Chemical Engineering, 14, 769-782.

Wolff, E. A., Skogestad, S., Hovd, M. and Mathisen, K. W. (1992). "A Procedure for Controllability Analysis", IFAC workshop on Interactions between Process Design and Process Control.

Zhao, Y. and Skogestad, S. (1997). "Comparison of Various Control Configuration for Continuous Bioreactors", Industrial and Engineering Chemistry Research, 36, 697-705.

Ziegler, J. G. and Nichols, N. B. (1943). "Process Lags in Automatic Control Circuits", Transactions of ASME, 65, 433.

Chapter 6

Consideration of General Disturbance Profiles

6.1 Introduction

This chapter explores the consideration of general disturbance profiles in process controllability assessments using the Dynamic Operability Framework. To accommodate this feature, the general disturbance profiles are sampled and the corresponding piecewise process dynamics are assessed within the framework. Furthermore, the dynamic semi-infinite MINLP algorithm of the framework is modified to enable sequential optimisation within a progressing window during the course of disturbances over the time horizon. This extension provides more adaptability and reduces conservatism of process design over longer time horizon and wider disturbance range and dynamics. The framework performance is demonstrated through a case study involving a nonlinear exothermic chemical reaction in a superstructure, which is affected by a set of disturbance profiles. The variations of process profitability, controllability, and computational costs due to different optimisation windows are presented.

The early results of synthesis problem in this chapter have been published in a conference paper titled "The Effect of General Disturbance Profiles to Process Controllability", presented at the Fourth Postgraduate Electrical Engineering Symposium, Australia, 2003. The complete version titled "Process Unfalsification within the Dynamic Operability Framework", has been submitted for presentation at the ESCAPE-14 Conference.

The structure of this chapter is as follows: Section 6.2 discusses the previous consideration of disturbances and uncertainties in Dynamic Operability Framework. Then section 6.3 puts it in perspective among various approaches addressing the effects of disturbances and uncertainties in process control. Based on this review, the framework expansion is outlined in section 6.3.6. It is followed by the formulation of the DOF process synthesis problem in section 6.4, and the implementation in section 6.5. Section 6.6 demonstrates the application of the proposed framework on the CSTR superstructure described in Chapter 5, and discusses the prospects of further improvements and expansions. Finally, section 6.7 summarises and concludes the chapter.

6.2 Disturbance and Uncertainty Consideration within the Dynamic Operability Framework

The previous chapter presented the process synthesis within the Dynamic Operability Framework (DOF) (Bahri, 1996), in which the Output Controllability Index (Vinson and Georgakis, 2000) has been incorporated. The variable redundancies associated with the index computation have been addressed by developing the redundancy analysis and elimination procedure based on the correlation PCA analysis of process steady state data. Since the redundancy is strongly linked with process and controller structures, the analysis and elimination procedure is integrated within the Branch and Bound based algorithm of the dynamic MINLP solution for the process synthesis problem.

It has been established that the framework addresses the effect of external disturbances and process uncertainties on process regulatory performances and profitability. So far, these disturbances and uncertainties are combined in one set of the variable θ. Their realisations are step functions, which magnitudes are assumed to be varying within lower and upper bound with uniform probability distributions and asymmetric nominal values. This assumption is the natural extension of the early steady state flexibility studies (Halemane and Grossmann, 1983), which assumed scalar values of all process variables within the feasible operating space. The framework then pursues the optimum parameters and structures of process and controllers, which gives the highest profit objective Φ, while keeping the process within the feasible operating conditions, despite the presence of the given disturbances and uncertainties realisations.

However, there are three shortcomings due to this assumption to the solution of the current framework. The first is the usage of step functions. The step changes in θ over an infinite time horizon are the well-known cause of the most severe regulatory dynamics. This interpretation is consistent both in linear and nonlinear cases. Therefore, this function is widely used to assess the worst-case conditions of the process. However, the corresponding designs over a wide time horizon are inevitably conservative, since the worst cases are of low probability in magnitude and frequency. Secondly, the profit objective Φ is calculated based on a set of assumed nominal values of θ, which does not necessarily represent the effects of θ dynamics. Thirdly, the effects of fast and slow, as well as measured and unmeasured components of θ are not distinguished.

More importantly, step changes, along with other regular characterisations such as impulse and sinusoidal, are only produced by tightly controlled generators. Therefore, the assumption based on these characterisations actually contradicts their own definition as uncontrollable parameters. In general, it is more difficult to model disturbance profiles from first principles than it is to model the process, and therefore, the disturbance dynamics are often more uncertain than process models (Gokcek *et al.*, 2000).

On the other hand, those regular characterisations along with linear process model assumptions provide well known analytic solutions and properties widely used for controllability assessments, such as first order response profiles, time constants, overshoots and such. Any general nonlinear behavior is eventually compared with these, within some tolerance on bounded deviations.

More importantly, the efficient solvers for these regular problems are widely available. Therefore, to accommodate general characterisation, control techniques such as Model Predictive Control (MPC) are known to sample the general profiles at a certain rate, such that the simple θ characterisation and linear process assumptions are sufficiently valid. The corresponding solutions are then combined to reconstruct the whole problem.

This study adopts the sampling approach in order to enable the assessment of general disturbance profiles to process profitability and controllability. This analysis focuses on the specific subset of θ, which is the measurable external disturbances profiles θ_e. These profiles are sampled and held over a sampling period. It is considered as a step excitation to the process, of which responses can be generated with the available dynamic solver. Combining the responses over the whole time horizon provides the piecewise process dynamics, which can be assessed using the Dynamic Operability Framework.

The conservatism attributed to the fixed time horizon assessment is addressed by performing the framework within a progressing window over the time horizon. The process profit over this window is calculated simultaneously with process dynamics to represent the actual profitability subject to disturbances, and its time-average is maximised within the framework.

The expansion of the Dynamic Operability Framework with these features are presented and demonstrated in sections 6.3.6-6.6. In the following section, the current approaches to disturbance and uncertainty characterisations in process control designs are reviewed to provide foundation for the expansions.

6.3 Disturbance and Uncertainty Consideration in Process Control Design

It has long been accepted that the effects of various sources of disturbances and uncertainties should be considered at process control design. Various approaches have been developed to solve this problem. These include the process designs based on stochastic programming methods, as well as the control synthesis based either on the robust, predictive, or unfalsification control methods. These approaches are briefly reviewed in the following sub-sections.

6.3.1 Formulation of Stochastic Programming Problems

In the mathematical programming approach, process design under uncertainty leads to the optimisation problems involving uncertainties either in the objective functions or in the constraints, which are referred to as the stochastic programming problem. This problem may express either the objective functions or the constraints in terms of some probabilistic representation, such as expected value or variances. The general formulation of this problem is as follows (Kim and Diwekar, 2002):

$$\min_{z} P_1[\Phi(\mathbf{z},\theta)] \tag{6.1}$$

$$\text{s.t.} \quad \mathbf{P_2}[\mathbf{h_i}(\mathbf{z},\theta)] \quad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{P_3}[\mathbf{g_j}(\mathbf{z},\theta)] \quad \mathbf{j} \in \mathbf{I}$$

$$\mathbf{z} \in \mathbf{Z}$$

$$\theta \in \Theta$$

Here, $\mathbf{z}$ is a vector of decision variables and θ is a vector of disturbances and uncertainties. The objective function $P_1[\Phi]$ is optimised subject to the model $\mathbf{P_2}[\mathbf{h_i}]$ and the operational constraints $\mathbf{P_3}[\mathbf{g_j}]$, where any or all of them are probabilistic functions.

For example, if P_1 is the expected value, the above optimisation problem becomes:

$$\min_{\mathbf{z}} P_1[\Phi(\mathbf{z},\boldsymbol{\theta})] = \min_{\mathbf{z}} E_{\boldsymbol{\theta}}[\Phi(\mathbf{z},\boldsymbol{\theta})] \tag{6.2}$$

where E_θ is the mathematical expectation with respect to $\boldsymbol{\theta}$. The optimal solution and optimal value of (6.2) are Φ^* and $\mathbf{z}^*$ respectively.

6.3.2 Stochastic Programming Solutions

The main challenge of stochastic programming in chemical process design is evaluating the effect of uncertain conditions to the whole design. The approaches to solve this problem include the multi-period, chance-constrained, and parametric formulations.

The multi-period or scenario-based formulations (Grossmann and Sargent, 1978) is one of the most widely used approaches. This approach transforms the original problem into a multi-period or multi-scenario deterministic problem associated with the realisations of $\boldsymbol{\theta}$. The later may take form of scalar values or simple functions such as steps, with magnitudes assumed to be randomly distributed.

The scenarios can be generated implicitly with the use of iterative two-level optimisation algorithms. The examples include the Dynamic Operability Framework and its predecessors (Bahri, 1996; Bandoni *et al.*, 1994; Ekawati and Bahri, 2001; Figueroa, 2000a; Figueroa, 2000b; Swaney and Grossman, 1985a; b). The first level solves the nominal design. Then, the second level assesses the feasibility of this design for all $\boldsymbol{\theta} \in \Theta$. If the design is infeasible for some combinations of $\boldsymbol{\theta}$, then the combinations causing the maximum constraint violation are identified, and appended to the set of critical combinations $\boldsymbol{\theta}^k$. This set becomes the additional periods in the subsequent outer-level. After solving the updated outer-level, the feasibility is evaluated in inner-level. This iteration continues until no constraint violation is found in the feasibility stage. The modified version described in Chapter 3 identifies most of the critical combinations in the first inner-level, therefore the algorithm terminates at the second iteration. So far, the uniform distribution of $\boldsymbol{\theta}$ is assumed.

Alternatively, the explicit approach a priori assumes a probability density function for $\boldsymbol{\theta}$, and approximates the corresponding probabilistic expressions based on the weighed discrete samples of $\boldsymbol{\theta}$. Therefore, the approach can be applied in one level of an optimisation algorithm. The discretisation is based on Gaussian quadrature method (Halemane and Grossmann, 1983; Pistikopoulos and Grossmann, 1988a; c), where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is discretised at quadrature points $\boldsymbol{\theta q}$, and the corresponding weighting values are determined. Then the problem is transformed to a multi-period optimisation problem. Each period corresponds to the constraints at each $\boldsymbol{\theta q}$, and the objective value Φ becomes the linear weighted combinations of functions of $\boldsymbol{\theta q}$. The choice of quadrature is unlimited. However, it is important to note that the number of $\boldsymbol{\theta q}$ combinations is the product of the quadrature with the number of disturbance and uncertain variables. Therefore, increasing the accuracy by adding more quadrature points leads to the exponential growth in the number of periods. The attempts to significantly reduce these combinations for steady state problems are reported in Novak and Kravanja (1999; 2003). Many of these studies assume scalar $\boldsymbol{\theta}$ with Gaussian probability distributions (Dimitriadis and Pistikopoulos, 1995; Papalexandri and Pistikopoulos, 1994a; b; Pistikopoulos and Grossmann, 1988a; b; c).

The alternative to the quadrature discretisation is to generate a fixed number of scenarios associated with the assumed joint probabilistic samples of $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ in the second level of the two–level algorithms. The scenarios are generated using either Monte Carlo Sampling (MCS), Latin Hypercube Sampling (LHS) (Liu and Sahinidis, 1996) or Hammersley Sequence Sampling (HSS) (Bernardo and Saraiva, 1998; Kim and Diwekar, 2002). This approach depends on the capability of the respective random number generators to produce the representative samples of $\boldsymbol{\theta}$, which is discussed in Kim and Diwekar (2002). Various probability distributions are accommodated while avoiding exponential growth of the number of periods over the number of uncertain parameters. However, numerous samples are still required to reasonably predict the objective function and constraints. The assumed $\boldsymbol{\theta}$ realisations are also scalar (Bernardo and Saraiva, 1998; Kim and Diwekar, 2002; Liu and Sahinidis, 1996).

One idea that is complementary to the multi-period approach is the chance-constrained formulation. This approach classifies the effect of $\boldsymbol{\theta}$ on the inequality constraints $\mathbf{g_j}$ as 'hard' and 'soft'. The hard constraints are those that cannot be violated under any circumstances, as in the case of safety limitations. The soft constraints, on the other hand, tolerate some violation, such as in the case of product specifications. In this approach, the stochastic problem is reduced into a deterministic one by enforcing the following form of constraints:

$$\mathbf{P}\left[\mathbf{g_j}(\mathbf{z},\boldsymbol{\theta}) \le 0\right] \ge \boldsymbol{\gamma} \tag{6.3}$$

Here, P[•] is the probability of satisfying a corresponding constraint. The hard and soft constraints are accommodated by assigning appropriate γ. The values for soft constraints commonly account for the desired probability region, such as γ=0.995 if the probability is within three times the value of standard deviations in a normal probability distribution.

This approach addresses the hard constraints similarly to the multi-period approach, and deals with soft constraints using the expectation or variance of the objective functions and constraints. While many studies do not distinguish between hard and soft constraints explicitly, their implementation of chance-constrained approach can be identified as follows (Samsatli *et al.*, 1998):

1. Expectation only: The optimisation of the expected value of objective function (Bhatia and Biegler, 1998; Mohideen *et al.*, 1996)

$$\min_{\mathbf{z}} \mathrm{E}(\boldsymbol{\Phi}(\mathbf{z},\boldsymbol{\theta})) \tag{6.4}$$

2. Worst case scenario: The optimisation of the objective function under the worst influence of disturbance and uncertainties (Bahri, 1996; Bandoni *et al.*, 1994; Nishida *et al.*, 1981; Raspanti *et al.*, 2000)

$$\min_{\mathbf{z}} \max_{\boldsymbol{\theta}} \mathrm{E}(\boldsymbol{\Phi}(\mathbf{z},\boldsymbol{\theta})) \tag{6.5}$$

3. Weighted mean and variance: The multi-objective optimisation of the expected value and variance of the objective function (Samsatli *et al.*, 1998)

$$\min_{\mathbf{z}} \alpha \mathrm{E}(\Phi(\mathbf{z},\boldsymbol{\theta})) + (1-\alpha)\,\mathrm{var}(\Phi(\mathbf{z},\boldsymbol{\theta})), \quad |\alpha| < 1 \tag{6.6}$$

4. Guaranteed variance : The optimisation of the expected value of objective function subject to the constrained variances (Samsatli *et al.*, 1998)

$$\min_{\mathbf{z}} \mathrm{E}(\Phi(\mathbf{z},\boldsymbol{\theta})), \text{ s.t.} \mathbf{var}_{\boldsymbol{\theta}}(\mathbf{g_j}(\mathbf{z},\boldsymbol{\theta}) \le 0) \le \boldsymbol{\gamma} \tag{6.7}$$

5. Guaranteed expected value: The optimisation of the variance of objective values subject to constrained expected value (Samsatli *et al.*, 1998)

$$\min_{\mathbf{z}} \mathrm{var}(\Phi(\mathbf{z},\boldsymbol{\theta})), \text{ s.t. } \mathbf{E}_{\boldsymbol{\theta}}(\mathbf{g_j}(\mathbf{z},\boldsymbol{\theta}) \le 0) \le \boldsymbol{\gamma} \tag{6.8}$$

6. Guaranteed one-sided constraint violation: The optimisation subject to one sided constraint violation (Samsatli *et al.*, 1998)

$$\min_{\mathbf{z}} \Phi(\mathbf{z},\boldsymbol{\theta}), \text{ s.t.} \mathrm{E}_{\boldsymbol{\theta}}(\alpha \mathbf{g}_{\mathbf{j}1}(\mathbf{z},\boldsymbol{\theta}) + (1-\alpha)\mathbf{g}_{\mathbf{j}2}(\mathbf{z},\boldsymbol{\theta}) \le 0) \ge \gamma, |\alpha| \le 1 \tag{6.9}$$

7. Expected one-sided constraint violation (Bernardo and Saraiva, 1998; Samsatli *et al.*, 1998)

$$\min_{\mathbf{z}} \mathrm{E}_{\boldsymbol{\theta}}(\alpha \mathbf{g}_{\mathbf{j}1}(\mathbf{z},\boldsymbol{\theta}) + (1-\alpha)\mathbf{g}_{\mathbf{j}2}(\mathbf{z},\boldsymbol{\theta}) \le 0) \ge \gamma, |\alpha| \le 1 \tag{6.10}$$

The multi-period and chance-constrained approaches yield a scalar objective value that guarantees the process specifications under uncertainties. Parametric programming (Acevedo and Pistikopoulos, 1996; Bansal *et al.*, 2002) on the other hand, analyses the optimum objective profile over a multi-dimensional uncertain parameter space. For a one-dimensional $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the process starts by finding the optimum solution Φ_1 and the corresponding point θ_1. At the next step, the new optimum solution point Φ_b and the corresponding point θ_b are sought within the interval $\theta_1 \le \theta \le \theta_{max}$. If θ_b exists, it divides $\boldsymbol{\Theta}$ into two subintervals $\theta_1 \le \theta \le \theta_b$ and $\theta_b \le \theta \le \theta_{max}$. This parametric problem is repeated at all subintervals until no new feasible solution can be found within an interval.

In many studies, the components of θ are assumed uncorrelated, which leads to the easy computations of joint probability distributions when many components exist. In some cases, however, some degree of correlation is possible. For example, between the temperature and the composition of liquor feed, where the composition might be affected by the temperature, while both are considered as disturbances to a certain process. Rooney and Biegler (2001) capture the nonlinear correlation between uncertain parameters in a confidence region using the least square likelihood-ratio test. Both the joint probability distributions and nonlinear confidence regions of θ can be used in conjunction with the stochastic programming methods described earlier.

Another approach addressing the worst-case scenario is using a constraint aggregation function or smoothing functions (Raspanti *et al.*, 2000; Rooney and Biegler, 2003). The constraint aggregation function overestimates the set of inequalities, which originally defines the feasible operating conditions, with a single evolving curve. Therefore, the flexibility problem in the two-level stochastic programming approach becomes a single constraint problem. The smoothing function is a class of parametric smoothing function obtained by twice integrating the joint probability density function or the nonlinear confidence region of θ. It further reformulates the two-level problem into a single level NLP problem. Both constraint aggregation and smoothing functions are continuous, but complicated and not necessarily differentiable. The corresponding NLP is difficult to solve, despite the reduced number of optimisation levels (Raspanti *et al.*, 2000).

The problems of process designs are commonly large due to non-linearity and dynamics. Accordingly, the control design subproblems are addressed in limited manners, in terms of structure, such as matching manipulated variables with controlled variables or selecting between several control schemes; or 'fine tuning' of the selected and pre-designed controller parameters. When the control design is addressed separately from the process, the effects of θ are addressed in various manners. Being intuitive and simple, the PID control method has been the dominant control strategy in process industries. However, its control law does not explicitly address the

information of process dynamics, such as deadtime or nonlinearity (Bequette, 1991), as well as the profiles of $\boldsymbol{\theta}$. Therefore, the next section focuses on several control methods that explicitly address the characterisation and identification of $\boldsymbol{\theta}$, namely, the robust control, predictive control and data-driven control methodologies.

6.3.3 Uncertainty Characterisations in Robust Control

The robust control method (Vinnicombe, 2001, and references therein; Zhou and Doyle, 1998) mainly considers locally linearised process models. The less understood process information, such as nonlinearity, parametric uncertainties, and external disturbance excitations, are grouped in one set of bounded uncertainties $\boldsymbol{\theta}$. The general schema is given in Figure 6.1. These uncertainties are considered unstructured when only their magnitudes or norms are known; or structured, if some parametric information is available, such as the bandwidth. The nature of these uncertainties can be multiplicative $\boldsymbol{\theta}_m$, which alters the process gain and phase; or additive $\boldsymbol{\theta}_a$, which causes offsets.

The aim of this approach is to find a set of off-line static feedback controller parameters $\mathbf{K}$ that minimise the H_2 or H_∞ norm of process dynamics with respect to the uncertainties.

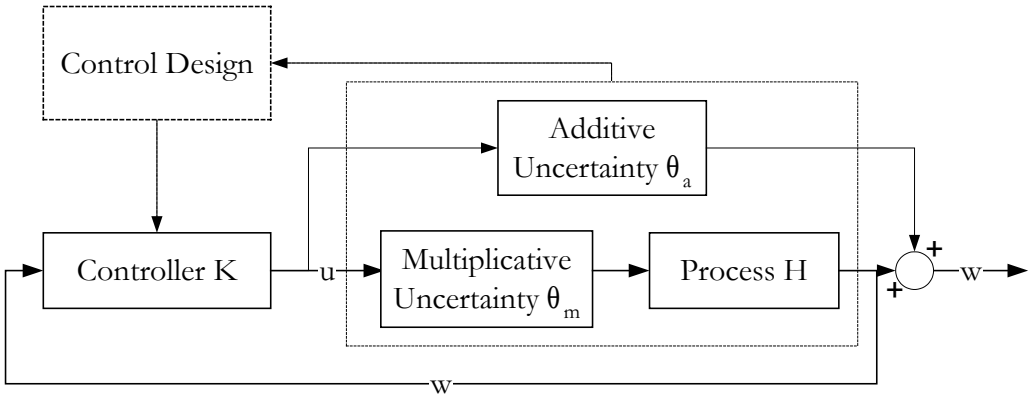


Figure 6.1 The Robust Control Scheme

The respective problems are represented in equations (6.11)-(6.12) as follows:

$$\min_{K} \|H\|_2 \tag{6.11}$$

$$\min_{K} \|H\|_\infty \tag{6.12}$$

Here, H is the process transfer-function from additive or multiplicative uncertainties, $\boldsymbol{\theta}_m$ or $\boldsymbol{\theta}_a$, to process outputs **w**. The controller structure, in terms of selection of manipulated and controlled variables are fixed in this approach.

Under the process linearity and unstructured uncertainties assumption, many studies of robust control characterise the uncertainties in frequency domain, in terms of their bandwidths and magnitudes. The corresponding control design includes shaping the outputs frequency spectrum (i.e. loop shaping controller) using a set of bandpass filters and static controller **K**. The time domain interpretations of unstructured uncertainty include white noise and persisting sinusoidal (Paganini, 1996). Therefore, minimising the norm of process performances subject to these uncertainties in the time domain is analogous to the unconstrained weighted variances (6.4) and worst-case (6.5) scenarios. Both ways yield conservative static robust controllers.

To reduce the conservatism, structural uncertainty approaches are proposed. These include assuming the uncertainties as colored noise, which is obtained from passing standard white noise through a linear finite-dimensional filter with uncertain coefficients. The results are the explicit formulas of output variances in terms of the uncertain parameters of the coloring filter (Davison *et al.*, 1999; 2000; Gokcek *et al.*, 2000).

The recent studies include expanding the H_∞ theory to nonlinear problems (Helton and James, 1999). Nevertheless, the approach does not provide clear connections between the uncertainty bounds or structure parameters to physical conditions of the process, and does not address process constraints.

6.3.4 Disturbance Models in Predictive Controls

The predictive control methods, which include the Dynamic Matrix Control (DMC), the Model Predictive Control (MPC) and the Generalised Predictive Control (GPC) approaches, have been used in process industries to provide on-line feed-back controllers, which may account for process constraints. The general scheme is given in Figure 6.2.

These approaches use the receding time concept, where at each sequence, the process impulse or step responses are predicted for R sequences, and the control action u is determined for L sequences. Only the first control sequence is applied until the next sequence is due. At the next sequence, the procedure is repeated to update the control action. Each control action and the corresponding controller parameters set K are the solution of a deterministic optimisation problem, which minimises a quadratic tracking or regulatory error subject to process constraints (Bemporad, 1998; Bequette, 1991).

The prediction phase uses the general discrete state space model as follows:

$$x(k+1) = A_p(x)x(k) + B_p(x)u(k) + B_d(x)\theta(k) + \xi_x(k) \tag{6.13}$$
$$w(k) = C_p(x)x(k) + D(x)u(k) + C_d(x)\theta(k) + \xi_x(k)$$
$$\theta(k+1) = \theta(k) + \xi_\theta(k)$$

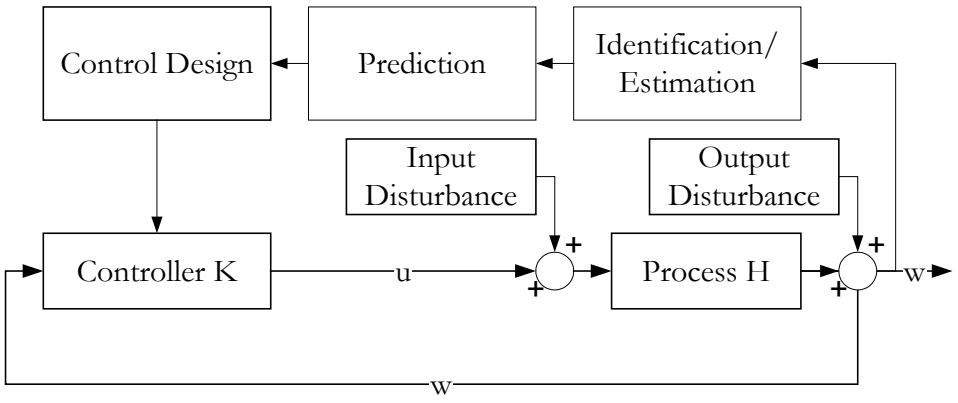


Figure 6.2 The Predictive Control Scheme

Here, x is the vector of state variables, u is vector of the manipulated variables, w is the vector of measured outputs, and θ is the vector of disturbances. Their values may vary at each sequence k. The vectors ξ_x, ξ_w, and ξ_θ are uncorrelated zero mean white noise random variables with covariances $Q_{\xi x}$, $Q_{\xi w}$ and $Q_{\xi\theta}$ respectively, representing the error between the actual and the predicted values.

The traditional predictive controls use the linear version of matrices A_p, B_p, B_d, C_p, D, and C_d. The disturbance vector θ may affect either the state or the output, where B_d and C_d are input and output disturbance model matrices respectively. These matrices are derived either by local linearisation of the original process models, by process identification, or estimation strategy (Bequette, 1991; Campbell and Rawlings, 1998). The corresponding process and controller structures are fixed. Recently, the problems have also been extended to nonlinear models (Bock *et al.*, 2000; Tenny *et al.*, 2002).

The general model (6.13) accommodates the augmentation of the disturbances to state variables. This is motivated by the recent trends in general model based controls, which augment the identified or estimated disturbances with the process states and use the information to design the controller (Bock *et al.*, 2000; Byrnes and Isidori, 2000; Chen *et al.*, 2000; Kadali and Huang, 2002; Kurtz and Henson, 1997; Liu and Peng, 2000; Marconi and Isidori, 2000; Marconi *et al.*, 2002; Pannocchia, 2003; Quintero and Quiennec, 2002; Tenny *et al.*, 2002).

Disturbance and state estimations in linear predictive controls commonly use the Kalman filter (Campbell and Rawlings, 1998; Krishnan and Hoo, 1999). This procedure principally minimises the error between the measured outputs and the linear model responses, ξ_x, ξ_w, and ξ_θ subject to $Q_{\xi x}$, $Q_{\xi w}$ and $Q_{\xi\theta}$, and accordingly estimates the states and disturbances. The Extended Kalman Filter (EKF) method is used for the receding horizon problems, where the measurements of fixed past sequence may be weighted relative to the current sequence (Quintero and Quiennec, 2002). Nonlinear disturbance estimation is recently reported in MPC and general nonlinear control design (Chen *et al.*, 2000; Liu and Peng, 2000; Tenny *et al.*, 2002).

A recent approaches for identification in predictive controls includes online identification of the sampled disturbance within a moving estimation window at each sequence (Rigopoulos *et al.*, 1997). The principal component analysis is repetitively applied to select the significant nodes, and accordingly the dynamic profiles are derived using Auto Regressive (AR) models.

The above procedures yield the input and output matrices of disturbance model, B_d and C_d, respectively. These notions clearly represent their points of entry. The output disturbance model ($C_d \neq 0$, $B_d = 0$) is frequently used in industrial applications for offset reduction (Tenny *et al.*, 2002). However, Pannocchia (2003) shows that the input disturbance models ($C_d = 0$, $B_d \neq 0$) provide rapid tracking and regulatory performances, while the combined input and output disturbance models ($C_d \neq 0$, $B_d \neq 0$) provide better offset reduction. Furthermore, the output disturbance model alone performs poorly subject to input uncertainty, in comparison to the input disturbance model.

These interpretations and utilisations of input and output disturbance models are similar to the structural multiplicative and additive uncertainties concept in robust control approach, where C_d and B_d determine the shape of the respective white noise filters. This addresses the connection to physical disturbances and the process constraints more explicitly in comparison to the robust control.

The use of EKF estimation or AR identification assumes the normality of disturbance data. This leads to the Linear Quadratic Gaussian (LQG) type of controller, which can be solved efficiently with the available quadratic optimisation solvers (Cervantes and Biegler, 1997; Leyffer, 2001; Tjoa and Biegler, 1991). The controllers are traditionally designed based on the open-loop predictions, subject to the worst-case disturbance. However, these lead to conservative control between the sequences, as well as limit the application to stable processes to guarantee bounded output responses. On the other hand, while closed-loop prediction leads to less conservative controller, it drops the linearity assumption, and increases the problem size with respect to the control horizon, as well as the computational costs (Bemporad, 1998).

6.3.5 General Disturbance Treatment in Data Driven Control

The data driven approaches include the iterative (windsurfing) adaptive control (Kosut, 2001 and references therein), the process and controller unfalsification (Safonov, 1996), and iterative identification and control (Veres, 1999 and references therein). The main idea is to design an on-line controller based on sampled closed-loop process data within a moving design window. The controller should be consistent with the sampled data, while satisfying the given specifications. The scheme is given in Figure 6.3.

This approach operates through two main procedures at different rates. The process control is sampled and held at fast sampling rate, which involves the inputs u, outputs w, and the controller parameter set K. The identification, estimation and control design are performed at a slower rate. These latter procedures are performed on the completion of L data sampling. Here, the interval is called design sequence. Between design sequences, the process uses the controller parameters determined in the previous sequence, which are based on the corresponding L sampled data. Therefore, the controller parameters are piecewise constants between design sequences.

The approach tries to avoid any assumptions about the process dynamics, as well as the disturbances and uncertainties. The corresponding control design uses the 'unfalsification' paradigm (Kosut, 2001; Safonov and Tsao, 1997; Woodley *et al.*, 1999).

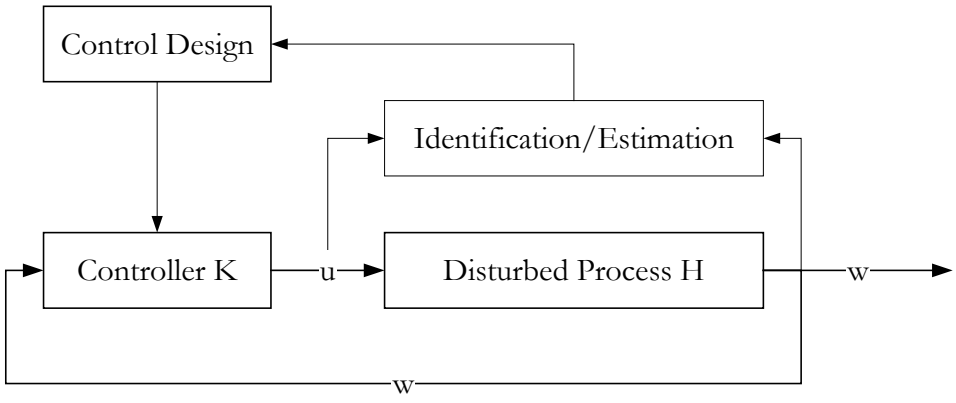


Figure 6.3 The Data Driven Control Scheme

Suppose **w** is the output vector corresponding to manipulation data **u** and affecting disturbances **θ**, where **W**, **U** and **Γ** are their respective functional spaces. The measurement set **M** contains L samples of input-output data $\mathbf{w_s}$ and $\mathbf{u_s}$ at the i^{th} collection sequence.

$$\mathbf{M} = \{(\mathbf{w},\mathbf{u}) \in \mathbf{W} \times \mathbf{U} \mid (\mathbf{w}_s, \mathbf{u}_s) \in (\mathbf{w},\mathbf{u})\} \tag{6.14}$$

$$\mathbf{P} = \{(\boldsymbol{\theta},\mathbf{w},\mathbf{u}) \in \boldsymbol{\Theta} \times \mathbf{W} \times \mathbf{U} \mid (\mathbf{w}_s, \mathbf{u}_s) \in \mathbf{M}\} \tag{6.15}$$

The control unfalsification is the selection of controller parameters **K** from a set of admissible parameters $\mathbf{K_a}$ such that the closed-loop specification **Φ** based on the measurement data **M** is close enough to the specification Φ_s. The quality of the 'closeness' is represented by Δ, which norm is bounded with a scalar value β. These are formulated in (6.16)-(6.17).

$$\mathbf{K} = \{\mathbf{K} \subset \mathbf{K_a} \mid \mathbf{u} = \mathbf{K}(\boldsymbol{\theta},\mathbf{w}), (\boldsymbol{\theta},\mathbf{w},\mathbf{u}) \in \mathbf{P}\} \tag{6.16}$$

$$\Phi(\beta) = \{\Phi \subset \Phi_s \mid \Phi(\boldsymbol{\theta},\mathbf{w},\mathbf{u}) - \Phi_s(\boldsymbol{\theta},\mathbf{w},\mathbf{u}) = \Delta, \quad \|\boldsymbol{\Delta}\|_\infty \le \beta\} \tag{6.17}$$

Any controller parameters that do not satisfy (6.17) are falsified and discarded from further consideration. Due to its data-driven nature, the whole unfalsification process is a large-scale optimisation problem.

This method implements the controller parameters based on the process performances during the previous sequence. Kosut (2001) shows that the unfalsified set **K** can only get smaller as one or both design sequences and sampling rates increase. There is no guarantee that the unfalsified closed-loop system will satisfy the desired performance specification at the future sequences, which underlines its adaptive inheritance.

The typical implementation of unfalsification controls is initialised within a finite set of candidate controllers. One arbitrary controller is implemented, and the closed-loop data is sampled. This data is used to falsify some of the candidate controllers with respect to a performance specification. Controller falsification (switching, updating) at each design sequence occurs if the performance level β is smaller than the measured performance **Φ**s. In other words, the external disturbances and uncertainties at each

design sequence must be sufficiently rich to trigger the falsification. The process is repeated until no further improvement is achieved.(Kosut, 2001; Woodley *et al.*, 1999).

This approach uses practically no assumptions on process model and disturbance characterisation in comparison to robust and predictive control. In other words, it accommodates general process and disturbance dynamics. The corresponding unfalsified controller is selected based on the sampled data, which leads to the least conservative controller for each design sequence. This nevertheless requires a large scale and expensive optimisation scheme. Therefore, so far this scheme uses the linear and static control parameterisation technique. In terms of structure, the switching between several different controller parameters for fixed manipulated and controlled variables have been reported (Brozenec *et al.*, 2001; Jun and Safonov, 1999; Safonov, 1996; Safonov and Tsao, 1997; Veres, 1999; Woodley *et al.*, 1999).

6.3.6 Generalisation of Disturbance Characterisation and Reduction of Design Conservatism

In comparison with the reviewed methods in section 6.3, the Dynamic Operability Framework developed so far (Bahri *et al.*, 1996, this thesis, Chapter 3-5) yields conservative process control design. This is mainly caused by characterising the disturbances and uncertainties as step changes, in which magnitudes are uniformly distributed between lower and upper bounds. Only by considering another alternatives to this worst-case characterisation, may the conservatism be reduced.

In the stochastic programming based design and robust control approaches, disturbance and uncertainty information has been gathered separately, which yields a fixed design over a time horizon. In contrast, the predictive and data-driven approaches use the latest sampled information to sequentially update the existing designs. The latter approaches provide opportunity to accommodate general measurable dynamic profiles, which are therefore adopted in this framework. Consequently, this limits the controllability analysis on process responses subject to the

measurable disturbances. To keep the problem to a manageable size, this study represents the process by mathematical model derived from the first principles.

The disturbance profiles are sampled and held over a sampling period, and the respective step responses are solved using the available DAE solver. The responses over the whole time horizon are combined to construct the piecewise process dynamics and assessed within the framework. The associated profitability is represented by maximising the time average of piecewise profit profile.

To further address the conservatism attributed to the long-term assessment, the problem is broken into smaller time intervals, resulting in a sequential framework. The adaptive approach is used for the closed-loop assessment of each sequence, for which the previously sampled data are used.

The dynamic MINLP feature of the framework naturally facilitates the unfalsification concept. It should be noted that this concept chooses the first-found feasible design and removes any prior falsified designs as the assessment progresses along the time horizon. This might end up with conservative design for the remaining sequences. Therefore, this study maintains the assessment of all design candidates at every sequence to explore all possible variations in designs over time, and keeps the original falsification as an optional feature.

The computational cost due to the computation of piecewise dynamics as well as process and controller structure problem is expected to be high. Furthermore, the adaptive feature does not guarantee convergence at each sequence. Instead, it indicates the requirement to alter the process specification at the respective operating region. The variations in computational cost and process feasibility are also demonstrated in the case studies. The modified framework formulation is presented in the next section.

6.4 Sequential Formulation of Dynamic Operability Framework

The Dynamic Operability Framework is modified to accommodate the general disturbance profiles and capture their effects to the objective function, as well as to allow sequential two-level optimisation over the time horizon. The modified formulation for each sequence is as follows:

For:

$$\begin{aligned}
&ws = 0, 1, \ldots, \text{round}((t_f - t_o)/\Delta t) - 1 \\
&op \in \{op : 1 \le op \le ws + 1\} \\
&\Delta t \in \{\Delta t : t_o \le \Delta t \le t_f\}
\end{aligned} \tag{6.18}$$

Outer-level:

$$\min_{\bar{\mathbf{z}}} \frac{\int_{t_{ow}-op\Delta t}^{t_{ow}} \Phi_{ws}(\bar{\mathbf{z}}, \boldsymbol{\theta}_e(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)\, dt}{op\Delta t} \tag{6.19}$$

$$\text{s.t.} \quad \mathbf{h_i}(\bar{\mathbf{z}}, \boldsymbol{\theta}_e(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t_{ow}-op\Delta t \le t \le t_{ow}} = 0 \qquad \mathbf{i} \in \mathbf{E}$$

$$\mathbf{g_j}(\bar{\mathbf{z}}, \boldsymbol{\theta}_e(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t_{ow}-op\Delta t \le t \le t_{ow}} \le 0 \qquad \mathbf{j} \in \mathbf{I_r}$$

$$\mathbf{g_{jd}}(\mathbf{y}) \le 0 \qquad \mathbf{jd} \in \mathbf{j}$$

$$\mathbf{r-OCI} = \frac{\mathbf{AOS}_{\boldsymbol{\theta}}(\bar{\mathbf{w}}(t))_{t_{ow}-op\Delta t \le t \le t_{ow}} \cap \mathbf{DOS}(\bar{\mathbf{w}})}{\mathbf{DOS}(\bar{\mathbf{w}})} = 1$$

$$\bar{\mathbf{z}} \in \{\mathbf{z}, \mathbf{y}, \mathbf{x_{ss}}, \mathbf{u_{ss}}, \mathbf{w_{ss}}\}$$

$$\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} : \mathbf{z^l} \le \mathbf{z} \le \mathbf{z^u}\}$$

$$\mathbf{y} \in \{0,1\}$$

$$\boldsymbol{\theta}_e(t)_{t_{ow}-op\Delta t \le t \le t_{ow}} \in \mathbf{EDS} = \{\boldsymbol{\theta}_e : \boldsymbol{\theta}_e{}^l \le \boldsymbol{\theta}_e \le \boldsymbol{\theta}_e{}^u\}$$

$$\bar{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}_{t_{ow}-op\Delta t \le t \le t_{ow}}$$

$$\mathbf{w} \in \mathbf{W} = \{\mathbf{w} : \mathbf{w^l} \le \mathbf{w} \le \mathbf{w^u}\}$$

$$\mathbf{x_{ss}} = \mathbf{x}(t = t_{ow} - op\Delta t)$$

$$\mathbf{u_{ss}} = \mathbf{u}(t = t_{ow} - op\Delta t)$$

$$\mathbf{w_{ss}} = \mathbf{w}(t = t_{ow} - op\Delta t)$$

$$t \in \{t : \max(0, t_{ow} - op\Delta t) \le t \le t_{ow}\}$$

$$t_{ow} = ws\Delta t$$

Inner-level:

$$\max_{\boldsymbol{\theta}} \mathbf{r} - \mathbf{OCI} = \frac{\mathbf{AOS}_{\boldsymbol{\theta}}(\overline{\mathbf{w}}(t))_{t_{ow}-op\Delta t\le t\le t_{ow}} \cap \mathbf{DOS}(\overline{\mathbf{w}})}{\mathbf{DOS}(\overline{\mathbf{w}})} \tag{6.20}$$

$$\text{s.t.} \quad \mathbf{h}_i(\overline{\mathbf{z}}^*, \boldsymbol{\theta}_e(t), \mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{w}(t), \mathbf{p}, t)_{t_{ow}-op\Delta t\le t\le t_{ow}} = 0 \quad \mathbf{i}\in\mathbf{E}$$

$$\overline{\mathbf{z}}^* \in \{\mathbf{z}^*, \mathbf{y}^*, \mathbf{x}_{ss}{}^*, \mathbf{u}_{ss}{}^*, \mathbf{w}_{ss}{}^*\}$$

$$\overline{\mathbf{w}}(t) \in \mathbf{w}(t) \in \{\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)\}_{t_{ow}-op\Delta t\le t\le t_{ow}}$$

$$\boldsymbol{\theta}_e(t)_{t_{ow}-op\Delta t\le t\le t_{ow}} \in \mathbf{EDS} = \{\boldsymbol{\theta}_e : \boldsymbol{\theta}_e{}^l \le \boldsymbol{\theta}_e \le \boldsymbol{\theta}_e{}^u\}$$

$$\mathbf{x_{ss}}^* = \mathbf{x}(t = t_{ow} - op\Delta t)$$

$$\mathbf{u_{ss}}^* = \mathbf{u}(t = t_{ow} - op\Delta t)$$

$$\mathbf{w_{ss}}^* = \mathbf{w}(t = t_{ow} - op\Delta t)$$

$$t \in \{t : \max(0, t_{ow} - op\Delta t) \le t \le t_{ow}\}$$

$$t_{ow} = (ws + 1)\Delta t$$

End.

The important feature of this formulation in comparison to the non-sequential version in Chapter 5 is the occurrence of various time scales. This framework involves several time units, which are illustrated in Figure 6.4. The framework progresses in **ws** sequences along the whole time horizon $t_o \le t \le t_f$. The period between optimisation sequences is Δt. Each sequence assesses the process controllability over an optimisation window $op\Delta t$, which is defined between $\max(0,t_{ow}-op\Delta t) \le t \le t_{ow}$. The disturbance profiles $\boldsymbol{\theta}_e(t)$ is sampled along this window, and held constant between sampling time t_s. These correspond to piecewise continuous output dynamics $\mathbf{w}(t)$, due to the piecewise constant profiles of $\boldsymbol{\theta}_e(t)$, $\max(0,t_{ow}-op\Delta t) \le t \le t_{ow}$.

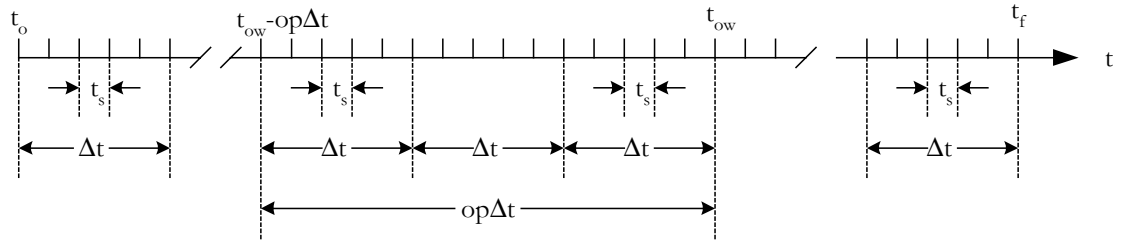


Figure 6.4 Time Scales within the Sequential Algorithm

Each component of $\boldsymbol{\theta}_e$ may have any form of dynamic profiles at different rates. For example, in the CSTR case, the dynamics of feed temperature may be much slower than that for the feed flowrate. Furthermore, the feed temperature profile tends to be sinusoidal, while the flowrate tends to be random. To capture these dynamics, the sampling rate should be frequent enough to reconstruct the fastest dynamics. A practical choice of t_s would be around 10% of the period of the fastest dynamics. One general dynamic profile for each $\boldsymbol{\theta}_e$ parameter is assumed over the time horizon, therefore only one set of profile combinations is considered in the sequential framework.

The iteration between the outer and the inner-levels of the framework is modified to assist the progression along the time horizon. The optimisation progresses one sequence at the inner-level, which means moves Δt forward in time horizon. The disturbance samples considered at each sequence are those within the optimisation window, which is the past opΔt period.

Since there is one sampled profile for each component of $\boldsymbol{\theta}_e$, the identification of the critical combination $\boldsymbol{\theta}^k$ is not required anymore. Instead, the inner-level computes the corresponding r-OCI and uses the value to trigger the subsequent outer-level. If r-OCI$\leq$1, then the outer-level is performed. Otherwise, the outer-level is skipped, and the subsequent inner-level is performed for the next sequence. The usage of r-OCI in this framework is analogous to the usage of the measured performance βs in the data-driven approach to enforce the process unfalsification in (6.17). This sequential algorithm is illustrated in Figure 6.5.

The dynamic profiles assessed in the optimisation window are piecewise continuous in respect to $\boldsymbol{\theta}_e(t)$. Accordingly, the AOS_θ contains the piecewise functional output profiles due to $\boldsymbol{\theta}_e(t)$. The controllability index r-OCI is calculated accordingly based on the intersection between AOS_θ and DOS. The objective profile $\Phi(t)$ is the subset of output profiles. Therefore, it is also a piecewise profile. Its maximum time average given in (6.19) is the optimum solution of the framework.

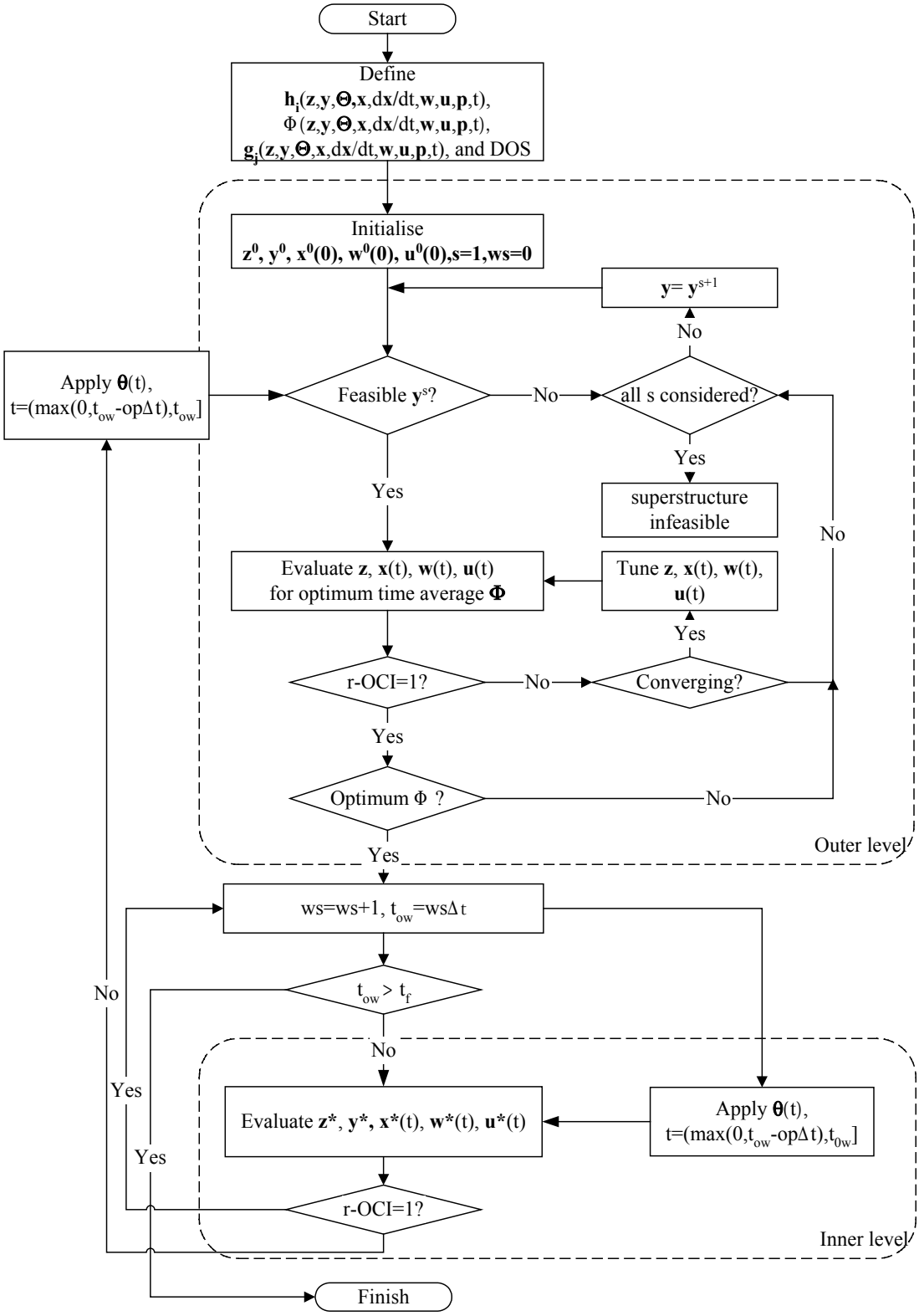


Figure 6.5 Sequential Optimisation Algorithm

The first outer-level initiates the optimisation sequence, where **ws** = 0, t_{ow} = 0, and there is no accumulated information about $\boldsymbol{\theta}_e(t)$. Therefore, this level performs the steady state MINLP at the nominal process condition similar to the non-sequential version. There is no calculation of AOS_{θ}, DOS, and r-OCI at this stage. The nominal optimum operating condition $\overline{\mathbf{z}}_{ss\,0}^{\,*}$ found in this level is sent to the inner-level for feasibility evaluation.

In the first inner-level, the sequence progresses so that **ws** = 1, and $t_o = \Delta t$. The disturbances $\boldsymbol{\theta}_e(t)$ are sampled over the optimisation window Δt, yielding the profile $\boldsymbol{\theta}_{e1}(t)$. Subsequently, the corresponding AOS_{θ}, DOS, and r-OCI are calculated based on the convex-hull of the dynamic profiles initialised at $\overline{\mathbf{z}}_{ss\,0}^{\,*}$. Since $\overline{\mathbf{z}}_{ss\,0}^{\,*}$ is typically located at the boundary of feasible operating region, the combined changes in $\boldsymbol{\theta}_{e1}(t)$ are likely to cause infeasible condition, which gives r-OCI<1, and triggers the unfalsification in the next outer-level.

The second outer-level solves the dynamic MINLP problem involving the calculation of AOS_{θ}, DOS, and r-OCI of the dynamic profiles due to $\boldsymbol{\theta}_{e1}(t)$. The condition of r-OCI=1 in this level guarantees the feasibility of the optimum solution of the first sequence $\overline{\mathbf{z}}_{ss\,1}^{\,*}$ subject to $\boldsymbol{\theta}_{e1}(t)$.

The second inner-level implies the progress to the second sequence, hence **ws** = 2, $t_{ow} = 2\Delta t$. If the final time t_f is reached, which is $t_{ow} \geq t_f$, then the assessment is concluded. Otherwise, feasibility test for the second sequence commences. The optimisation window becomes $2\Delta t$, giving the sampled profile $\boldsymbol{\theta}_{e2}(t)$. If r-OCI<1 at this level, the unfalsification is required and the next outer-level is performed subject to $\boldsymbol{\theta}_{e2}(t)$. Otherwise, the second outer-level is skipped over to the inner-level for the next sequence. This sequence continues until the whole time horizon is covered.

The size of the optimisation window, opΔt, may vary between Δt and t_f. The range definition $t \in \{t : \max(0, t_{ow} - op\Delta t) \le t \le t_{ow}\}$ indicates that this window builds up during the first **op** sequences. Once the full window is built, it progresses along the time horizon. At constant disturbance sampling rate and sufficiently rich disturbance variation between sampling, the computational cost of piecewise output profiles is higher over a wider window. On the other hand, the design change associated with the unfalsification is smoother at wider windows. Therefore, some trade-off between the smoother design changes and computation cost of piecewise profiles should be considered.

From the second outer-level onwards, the sequential framework yields a delayed optimum closed-loop design for the corresponding optimisation window. Any probable major upset between two sequences is compensated by the control strategies selected at the earlier sequence. The latter sequence only updates the earlier strategy if the disturbance variation drives the process infeasible between the sequences. This does not guarantee complete feasibility at every sequence, an expense of the reduced design conservatism. Compared to the earlier version described in chapter 5, the computational complexity of this framework is one order higher by the number of sequences.

This framework maintains its dynamic optimisation formulation, where the initial values are selected such that the optimum objective is achieved as long as the corresponding dynamics are feasible. This is principally different from the adaptive dynamic optimisation, where the initial conditions are constrained to the existing values, and the dynamic profiles are constrained to achieve the optimum objective. Acquiring this adaptive dynamic optimisation approach requires drastic change in problem formulation and optimisation techniques. Therefore, this is left for future works.

6.5 Framework Implementation

In this study, the analysis of the effects of general disturbance signal is based on simulated data. The major numerical modifications in this sequential framework are the samplings of general disturbance profiles and the computation of the respective piecewise output profiles. The full falsification feature, which discards the falsified design candidates as the assessment moves along the time horizon is also discussed.

To accommodate different profiles of each disturbance parameters, all profiles are sampled and held within sampling period t_s. In this study, this task is implemented using the functions *gendistrand2.m* and *simple2.m*, which are included in the DOF toolbox.

The function *gendistrand2.m* accepts one or more sets of disturbance profiles, where each set contains an array of time and the corresponding instantaneous profile values. The time array may be irregular, but must be monotonically increasing. When there are two or more disturbance profiles, the profiles are combined into a matrix, where the first column is the combined time sampling, and the other columns represent each of the corresponding disturbance profiles. The combining operation producing this matrix is such that one row corresponds to one change in one profile at a time. If two or more changes occur at one instantaneous time, two or more rows are required to represent them. The function *simple2.m* is subsequently used to combine these rows to shorten the matrix as much as possible. These functions are organised in disturbance sampling algorithm illustrated in Figure 6.6.

The sampling of two sinusoidal disturbance profiles is illustrated in Figure 6.7. The original profiles have different frequencies and different sampling rates. The *gendistrand2.m* and *simple2.m* combines and holds these profiles over the common sampling period.

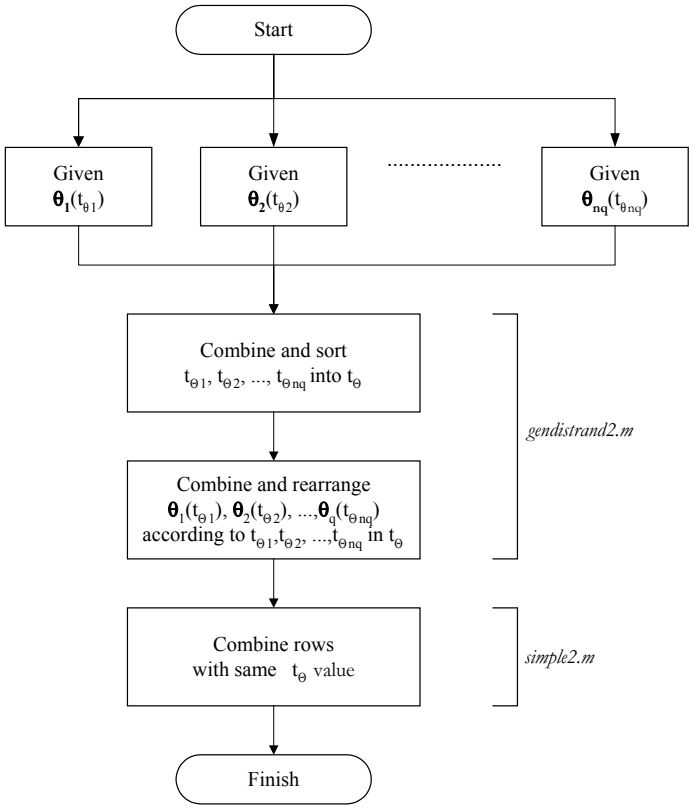


Figure 6.6 Disturbance Sampling Algorithm

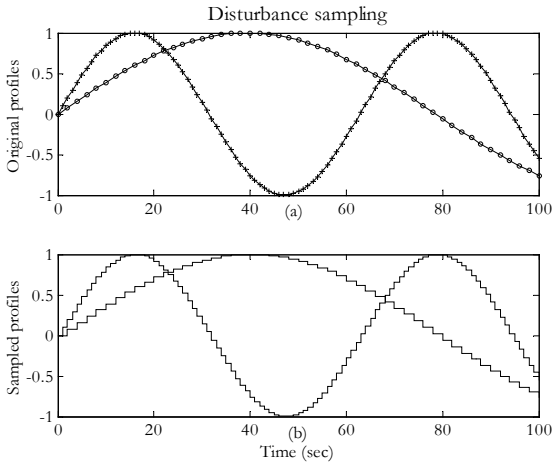


Figure 6.7 Disturbance Sampling
(a) Original Profiles (b) Sampled Profiles

To generate the piecewise output profiles, the corresponding DAE is solved consecutively for each sampling period. The dynamic values at each end of a period become the initial conditions for the next one, yielding piecewise continuous dynamic profiles. The template to produce this piecewise DAE solution is given in *seqdae1.m* and *seqdae2.m*, and their algorithms are illustrated in Figures 6.8 and 6.9.

The inputs for *seqdae1.m* are the optimum design solution from the preceding outer-level, $\overline{\mathbf{z}}_{i-1}^{*}$, where $i = 1, 2, \ldots, ws$; the sampled disturbance profiles, and the combined sampling time. The outputs are the sampling time, the piecewise output and the affecting disturbance profiles. The objective profile $\mathbf{\Phi}(t)$ is the subset of the output profiles $\mathbf{w}(t)$, therefore it is also piecewise continuous. The respective time average is computed by integrating its piecewise solution over the optimisation window, then dividing the result with the size of the optimisation window. This DAE solver template is used within the framework.

The *seqdae2.m* is used to simulate the process responses due to the sequential design solutions of this framework $\{\overline{\mathbf{z}}^{*}_{1}, \overline{\mathbf{z}}^{*}_{2}, \ldots, \overline{\mathbf{z}}^{*}_{ws}\}$ in addition to the disturbances, once the sequential controllability assessment is completed. The changes in design, if any, are considered at the beginning of the time sequence t_{ow}, while the disturbances are considered at every time sampling t_s. The simulation would show the changes in design decision during the course of disturbances, as would be demonstrated in Section 6.6.

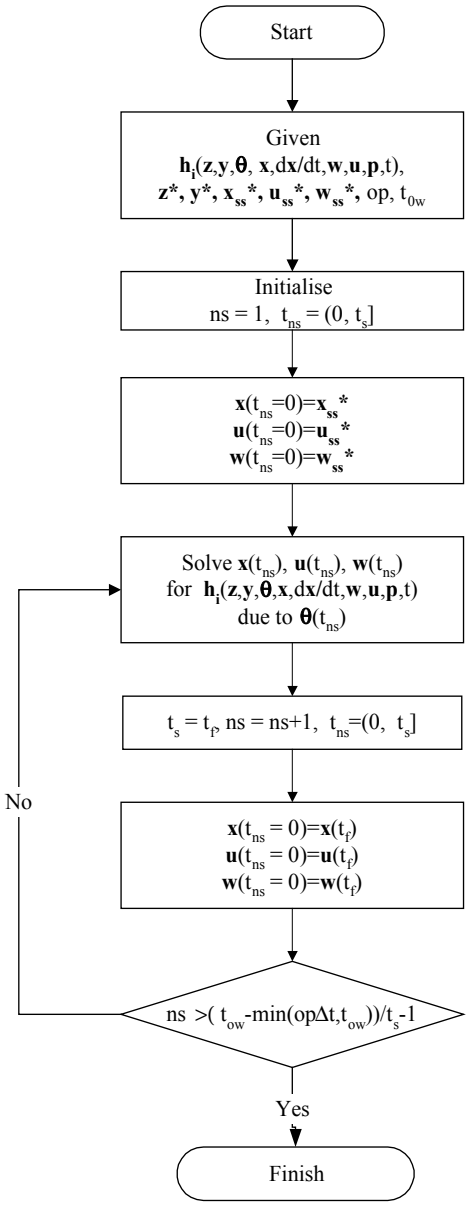


Figure 6.8 *Seqdae1.m* Algorithm

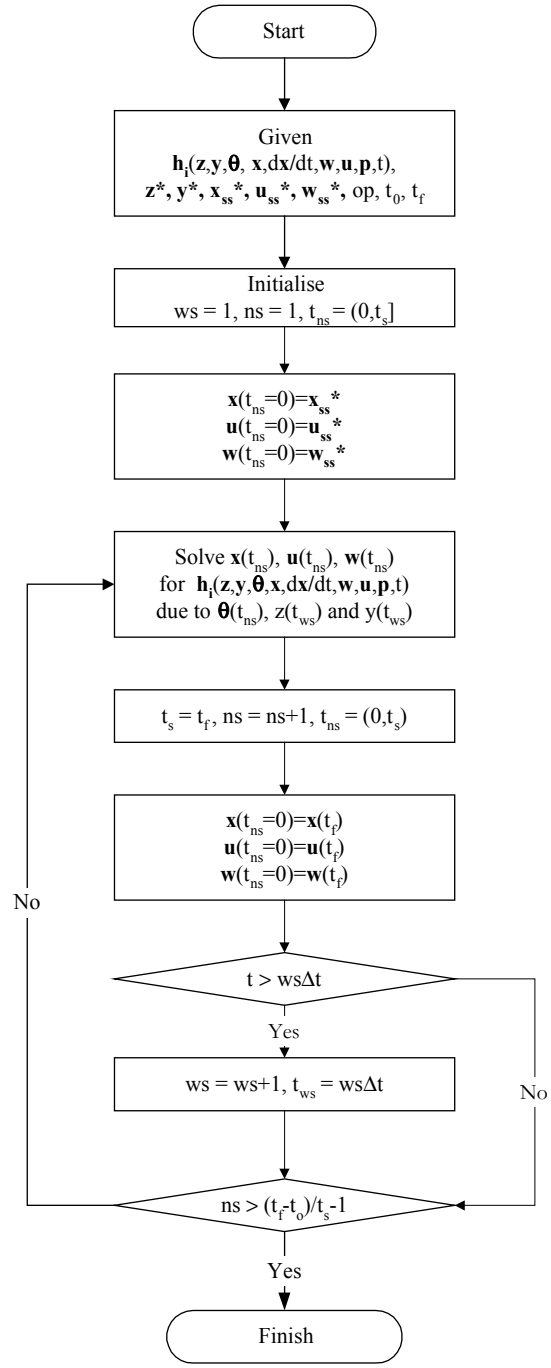


Figure 6.9 *Seqdae2.m* Algorithm

Both inner and outer-levels in this sequential framework consider one set of piecewise dynamic profiles over an optimisation window. These data are mostly transient; even if taken from the end of each sampling time. This reduces the collinearity among redundant variables, and therefore requires a higher collinearity tolerance at the redundancy analysis and elimination procedure within this framework. In this sequential framework, 0.1 tolerance is used, instead of 0.05 in the non-sequential version.

The falsification of the design parameter and structure in this framework is facilitated by the dynamic MINLP nature of the framework. In this framework, the optimum design at each sequence is pursued. Alternatively, the complete accordance to falsification concept can also be realised at the outer-levels, by updating the range and combinations of the continuous and binary variables, and terminating the Branch and Bound algorithm once the first feasible structure is found. These features are kept optional at this sequential framework.

Either way, the sequences of optimum profit, as well as process and controller structures and parameters, are generated. The variation between each sequence as well as the computational costs depend on the size of the optimisation window, which are demonstrated further in the following case studies.

6.6 Case Studies

The following case studies are the expansion of the process and controller synthesis of the CSTR superstructure described previously in section 5.4.4. The process and controller equations have been given in (5.29)-(5.51).

The process is assessed subject to the dynamics of the feed composition $C_f(t)$ and the feed temperature $T_f(t)$. In this study, these profiles are simulated over the time horizon. While there are infinite possible profiles, three particular cases are considered:

1. Exponential profiles
2. Sinusoidal profiles
3. Random profiles

The exponential profiles represent the case where significant dynamics are concentrated in a certain period. The other two represent the case where the dynamics are distributed evenly, in regular manner at sinusoidal profiles, and irregularly at random profiles. A general disturbance profile may contain the combination of any of the above profiles at some degrees. These cases are presented to provide gradual analysis from the most familiar results to the least known.

There are four conditions evaluated within 500 seconds time horizon:

1. Time sequence 50s, time window 50s
2. Time sequence 50s, time window 150s
3. Time sequence 50s, time window 300s
4. Time sequence 500s, time window 500s

These different time windows are used to evaluate the computational costs of solving the piecewise dynamics within the framework. The 500s window provides the non-sequential optimisation over the whole time horizon and the benchmark solution for the other windows. The results are also compared with the results from non-sequential framework presented in Chapter 5.

In these case studies, the focus is in assessing whether the optimum design can change during the course of disturbances, rather than discarding the falsified candidates. Therefore, the falsification feature is kept off. These case studies are performed using a Pentium IV PC 2.4 GHz, with 256 MB RAM.

6.6.1 Exponential Disturbance Profiles

The mathematical expression of the disturbance profiles is as follows:

$$\theta_e(t) = \begin{cases} T_f(t) = \begin{cases} 300 & 0 \le t < 25 \\ 300 + 15e^{-t/40} & 25 \le t < 500 \end{cases} \\ C_f(t) = \begin{cases} 20 & 0 \le t < 25 \\ 20 + e^{-t/60} & 25 \le t < 500 \end{cases} \end{cases} \tag{6.21}$$

Both profiles are provided at 25-second intervals. Therefore, the samples are also provided at the same intervals, as shown in Figure 6.10.

The 500s window assesses the given disturbance profile for the whole time horizon in one sequence. Therefore, it completes in two iterations. Since it has the whole information about disturbance profile, its solution becomes the benchmark for other windows. The 50s, 150s and 300s windows, on the other hand, are built during the first few 50s time sequences. Once completed, these windows progress until the final time is reached. The assessment completes in 11 iterations.

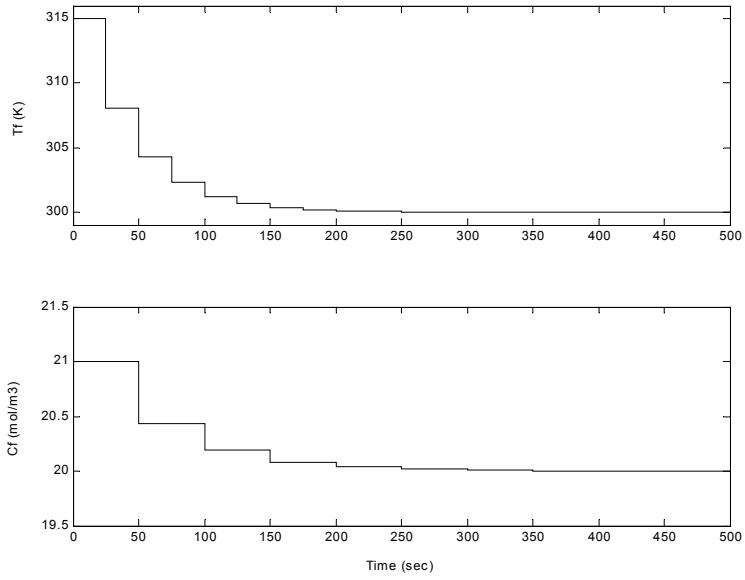


Figure 6.10 Exponential Disturbance Profile

Table 6.1 Controllability Assessment Results: Exponential Disturbance Profile

	Non-sequential	Sequential approach			
		opΔt=50s (average)	opΔt=150s (average)	opΔt=300s (average)	opΔt=500s
Φ \$/h	106.542	144.4015	144.3918	144.3855	144.3714
T_{fN} K	300	301.6138	301.6138	301.6138	301.6138
C_{fN} mol/m^3	20	20.1467	20.1467	20.1467	20.1467
Q_{f1} m^3 (y_{p1})	0.356 (1)	0.3389 (1)	0.3389 (1)	0.3389 (1)	0.3389 (1)
Q_{f2} m^3 (y_{p2})	0.288 (1)	0.2913 (1)	0.2913 (1)	0.2913 (1)	0.2913 (1)
Q_3 m^3 (y_{p3})	0.050 (1)	0.0000 (0)	0.0000 (0)	0.0000 (0)	0.0000 (0)
Q_4 m^3 (y_{p4})	0 306 (1)	0.3389 (1)	0.3389 (1)	0.3389 (1)	0.3389 (1)
α_1 (y_{c1})	2 (1)	2.0000 (1)	2.0000 (1)	2.0000 (1)	2.0000 (1)
α_2 (y_{c2})	2 (0)	2.0000 (0)	2.0000 (0)	2.0000 (0)	2.0000 (0)
V_1 m^3	4.500	4.5000	4.5000	4.5000	4.5000
V_2 m^3	4.500	4.5000	4.5000	4.5000	4.5000
CPU time (min)	5.023	2.1933	7.1348	11.0043	14.1313

Table 6.1 compares the closed-loop optimum solutions of the non-sequential with the sequential frameworks. The non-sequential framework finds a combined parallel and series process structure ([y_{p1}, y_{p2}, y_{p3}, y_{p4}] =[1, 1, 1, 1]), while the sequential framework finds purely parallel process structure ([y_{p1}, y_{p2}, y_{p3}, y_{p4}] =[1, 1, 0, 1]). Both frameworks find the same controller structure ([y_{c1}, y_{c2}] =[1 0]).

The table also shows the varying optimum profit Φ between the 50s, 150s and 300s windows. Process parameters and structure, represented by $Q_{f1\text{-}2}$, $Q_{3\text{-}4}$, $y_{p1\text{-}4}$, and $V_{1\text{-}2}$, do not change over the time horizon. On the contrary, controller structure and parameters represented by $y_{c1\text{-}2}$, $\alpha_{1\text{-}2}$, T_{1ss}, T_{2ss}, $Cool_{1ss}$, $Cool_{2ss}$, C_{pss}, do change over the time horizon. These changes are illustrated in Figures 6.11-6.12.

The different profit values Φ found by the non-sequential and the sequential framework are due to the assumed disturbance profiles, distribution and nominal values. The non-sequential framework assumes a uniform distribution of the step magnitudes, of which nominal values are significantly closer to the lower bounds. The respective profit therefore represents the nominal profit when the process is kept feasible despite the worst-case step changes of the disturbances. The significant disturbance dynamics in the sequential framework, in contrast, is concentrated at the first 50s and less severe afterwards. For most of the time, the disturbances are close to the nominal values. In addition, the framework optimises the profit time-average, not the nominal one. Therefore, the optimum profits found by the sequential framework are higher compared to the non-sequential version.

The computational costs in terms of CPU times are expectedly higher at the 500s window, because it involves 20 piecewise dynamics, which are repeatedly solved in the outer-level. Since the first 50s contain the significant dynamics, the optimum design found at the first window guarantees the closed-loop feasibility for the subsequent windows. Therefore, once the first window is completely built, no further falsification is required, and the remaining outer-levels are skipped. This explains the fast result obtained at 50s windows. The 150s and 300 s windows need two and five more sequences respectively to complete their windows before skipping the remaining outer-levels.

The applications of the design sequences to the process are shown in Figure 6.13. The profiles at the 500s window are completely feasible, because the whole profile is solved in one sequence. The other windows show the delayed compensation, because the disturbances occur at 25s, while the first control sequence is computed and applied after 50s. During the first 50s, the process are initialised at the optimum steady state values, which is easily driven infeasible, until the controller design is updated at the second sequence. Afterwards, the disturbance variations are completely manageable, therefore no further updating is required. Therefore, the design changes are only occurs at t = 50s for the 50, 150 and 300s windows, as shown in Figures 6.11 and 6.12.

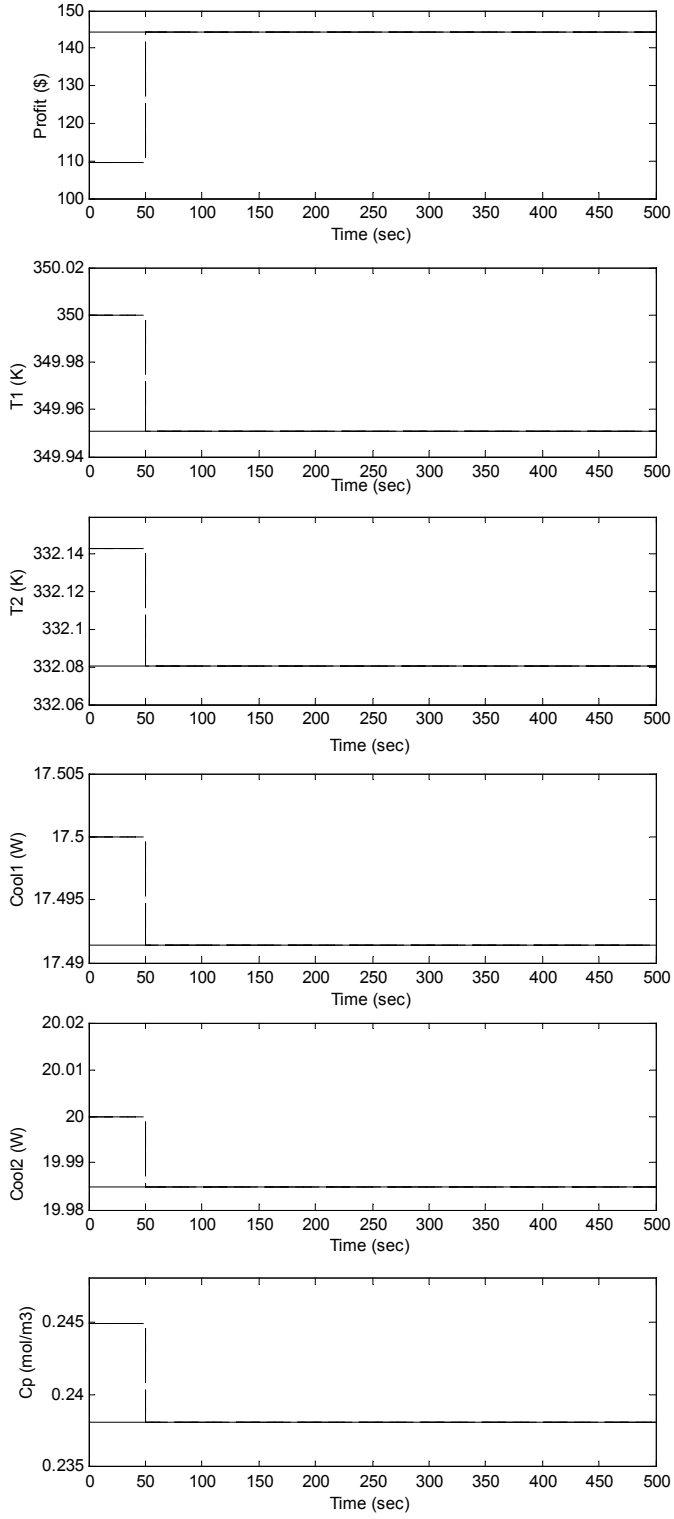


Figure 6.11 Profit and Initial Condition Sequences: Exponential Disturbance Profile

——— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

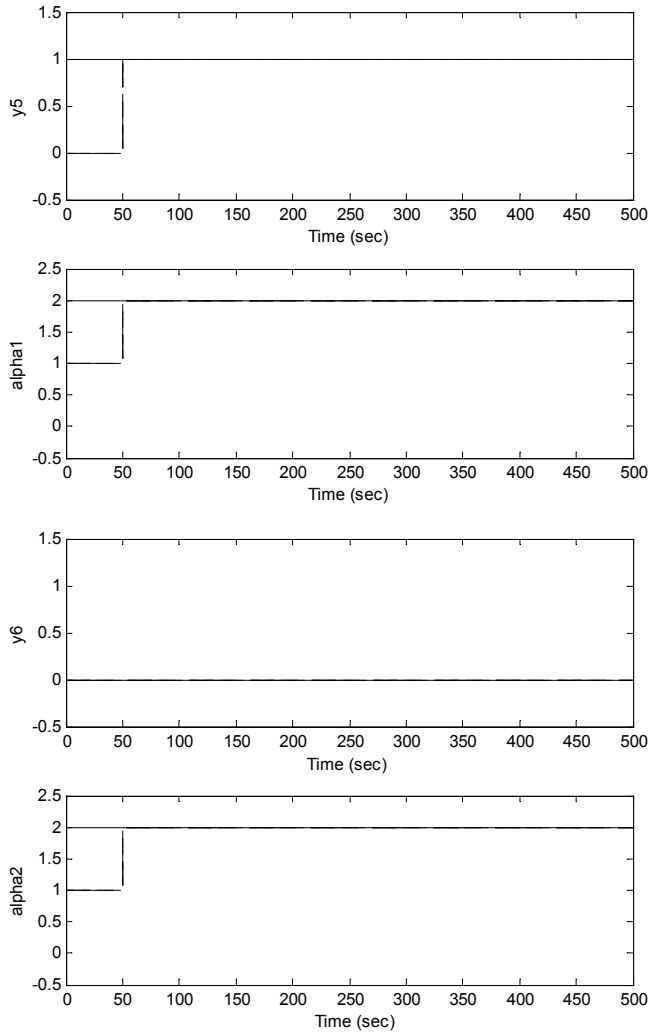


Figure 6.12 Controller Structure: Exponential Disturbance Profile

——— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

Figure 6.13 shows the process responses when the process design variation are applied to the process. The discontinuity of process responses at the sequence time t_{ow}, occurs because within the 50s, 150s and 300s windows, their initial values at each sequence are adjusted according to the optimum solution found in the previous sequence. Applying the design sequences to the process is analogous to introducing new manipulated values ($\alpha_{1\text{-}2}$) and new reference values (T_{1ss}, $Cool_{1ss}$, T_{2ss}, and $Cool_{2ss}$). This can drastically change the direction of disturbance compensation, resulting in the rough responses. Suggestion to reduce this problem is given in section 6.6.4.

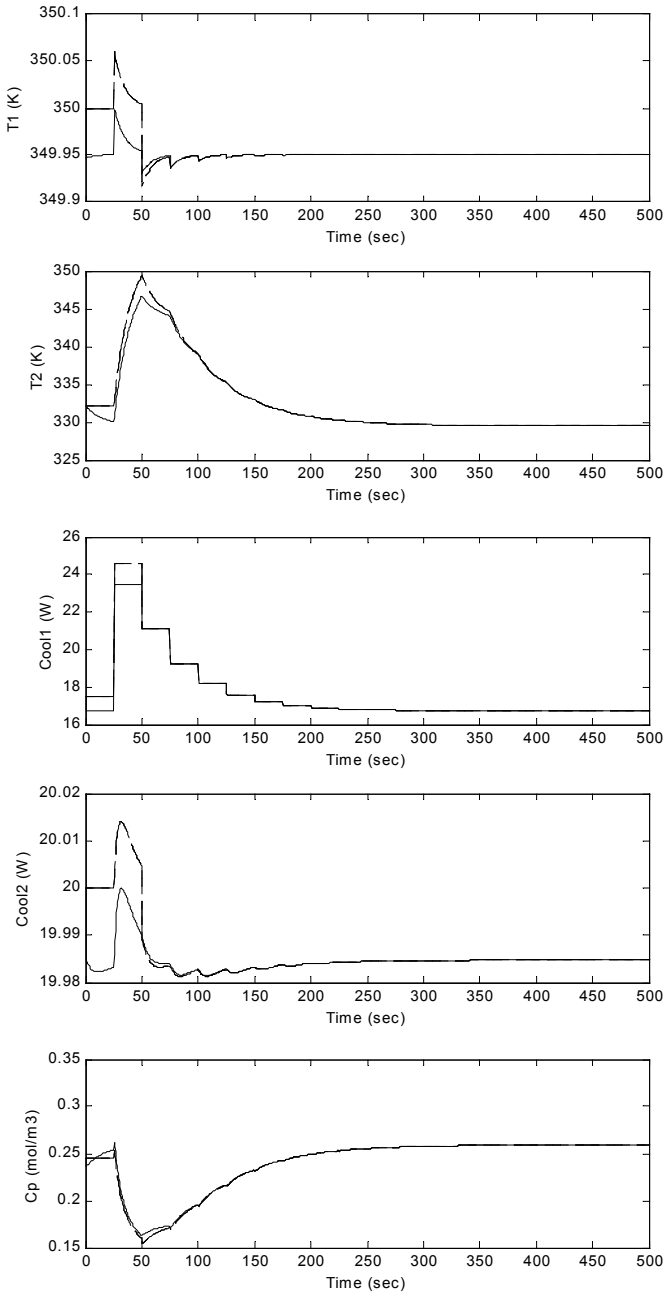


Figure 6.13 Overall Process Responses: Exponential Disturbance Profile

—— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

6.6.2 Sinusoidal Disturbance Profiles

The mathematical expression of the sinusoidal profiles is given in (6.22). The data for feed temperature is available every 25s, and 50s for the feed concentration. Hence, the combined sampling interval is 25s, and the profiles are shown in Figure 6.10.

$$\theta_e(t) = \begin{cases} T_f(t) = 306.5 + 8.5\sin\left(\dfrac{t-65}{50}\right) & 0 \le t < 500 \\ C_f(t) = 20.25 + 0.75\cos\left(\dfrac{t-110}{70}\right) & 0 \le t < 500 \end{cases} \tag{6.22}$$

The disturbance dynamic is spread evenly over the time horizon, with time averages $[T_{fN}, C_{fN}]$ = [307.6913, 20.2436]. This combination is higher than the nominal values used by the non-sequential version. However, their variation over time is such that if T_f goes up, C_f goes down, and the vice-versa. In the non-sequential assessment, these disturbance combinations are not critical. Therefore, the process can operate closer to the given constraints.

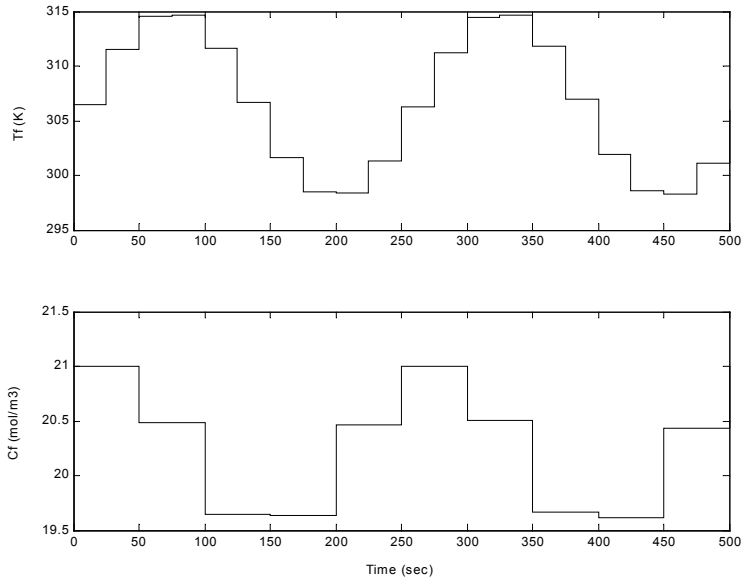


Figure 6.14 Random Disturbance Profile

Table 6.2 shows that the sequential assessments find higher time averaged profits due to the given profiles, in comparison to the non-sequential ones. Similar to the exponential case, the sequential framework finds a purely parallel process structure ($[y_{p1}\ y_{p2}\ y_{p3}\ y_{p4}] = [1\ 1\ 0\ 1]$) and controller structure ($[y_{c1}\ y_{c2}] = [1\ 0]$) at all cases.

The disturbance variation contained in every sequence and window is significantly richer than the exponential case. This contributes to high computational cost (in CPU time) of the sequential assessments. The cost of building a 300s window in six sequences is higher than the assessment of a 500s window in one sequence. This calls for smart selection of window and sequence sizes. The corresponding changes in design variables are shown in Figures 6.15 and 6.16.

Table 6.2 Controllability Assessment Results: Sinusoidal Disturbance Profile

	Non-sequential	Sequential approach			
		opΔt=50s (average)	opΔt=150s (average)	opΔt=300s (average)	opΔt=500s
Φ \$/h	106.5420	131.8034	131.6902	131.6893	131.6650
T_{fN} K	300	307.6913	307.6913	307.6913	307.6913
C_{fN} mol/m^3	20	20.2436	20.2436	20.2436	20.2436
Q_{f1} m^3 (y_{p1})	0.3561 (1)	0.3005 (1)	0.3004 (1)	0.3004 (1)	0.3004 (1)
Q_{f2} m^3 (y_{p2})	0.2882 (1)	0.2657 (1)	0.2655 (1)	0.2654 (1)	0.2653 (1)
Q_3 m^3 (y_{p3})	0.0500 (1)	0.0000 (0)	0.0000 (0)	0.0000 (0)	0.0000 (0)
Q_4 m^3 (y_{p4})	0 3061 (1)	0.3005 (1)	0.3004 (1)	0.3004 (1)	0.3004 (1)
α_1 (y_{c1})	2.0000 (1)	2.0000	2.0000	2.0000	2.0000
α_2 (y_{c2})	2.0000 (0)	2.0000	2.0000	2.0000	2.0000
V_1 m^3	4.5000	4.5000	4.5000	4.5000	4.5000
V_2 m^3	4.5000	4.5000	4.5000	4.5000	4.5000
CPU time (min)	5.0230	6.7090	12.3152	38.3315	16.8878

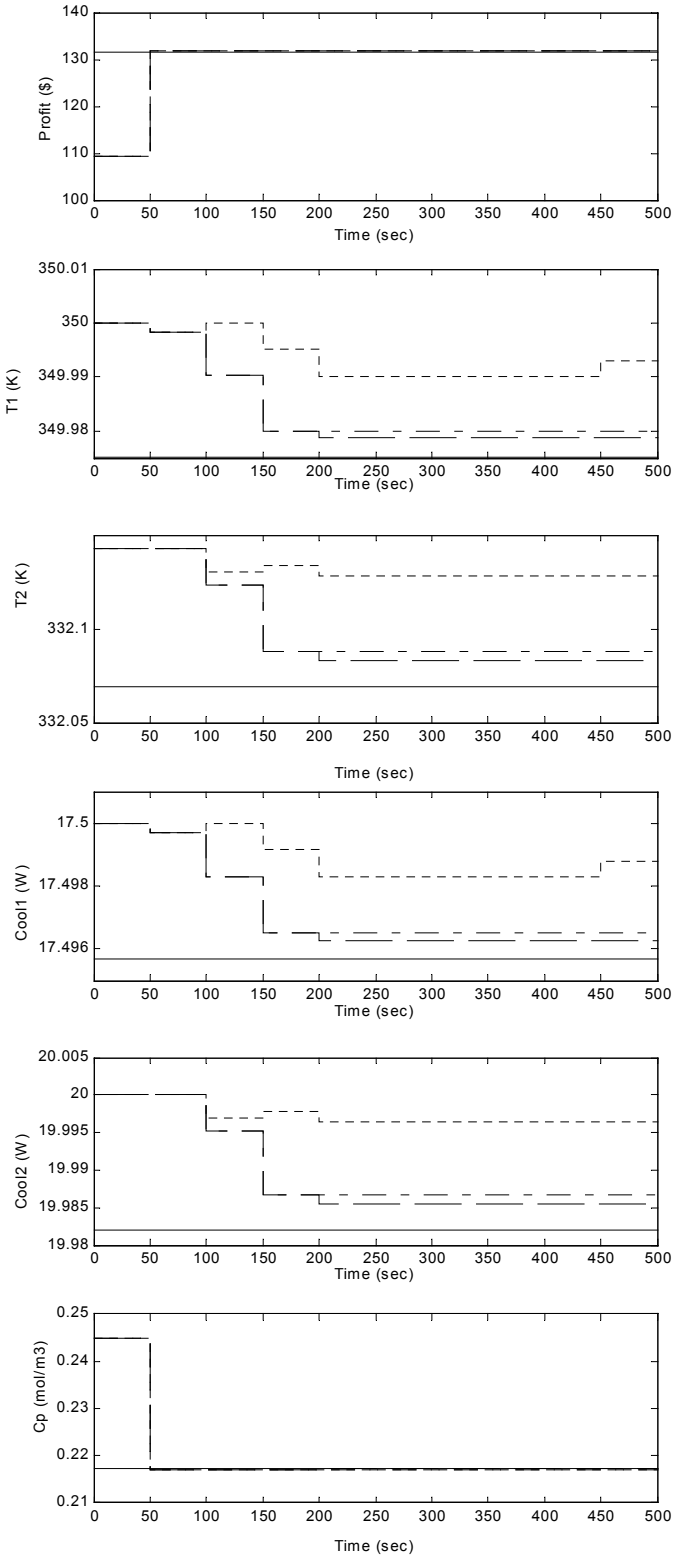


Figure 6.15 Profit and Initial Condition Sequences: Sinusoidal Disturbance Profile

——— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

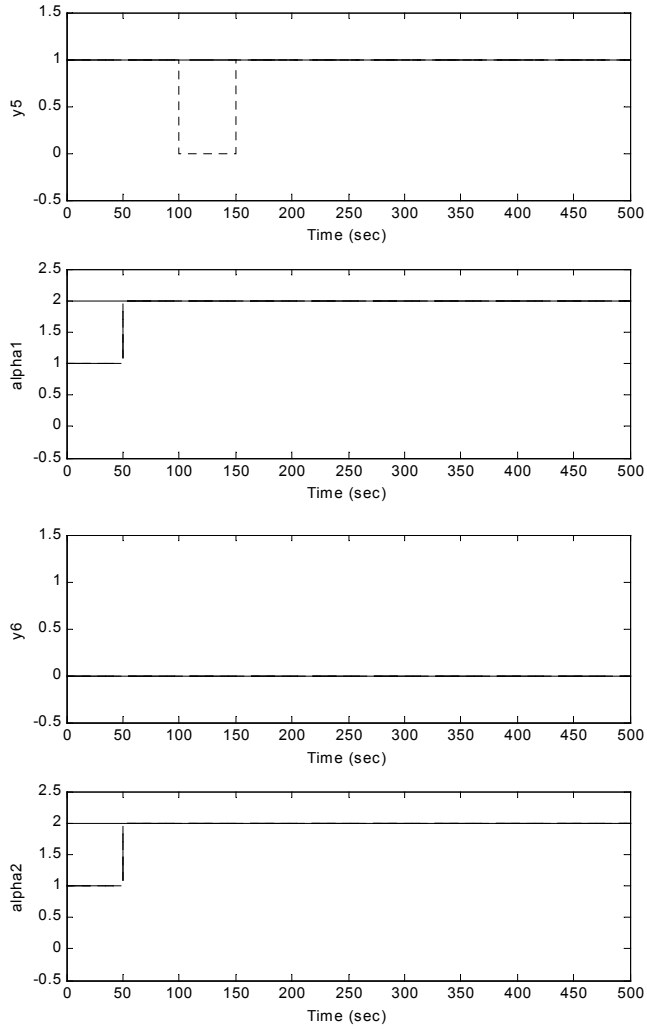


Figure 6.16 Controller Structure: Sinusoidal Disturbance Profile

——: opΔt=500, - - -: opΔt=50, — -: opΔt=150, ——: opΔt=300

Figures 6.15 and 6.16 show that over the sequences, the solutions of other windows tend to converge to the solution of 500s window. Figure 6.17 further shows this convergence when the design sequences are applied.

The 50s window is apparently sensitive to the disturbance. This is shown by the controller structure switching at the first reactor from m_{c1}-T_1 (y_{c1}=1) to m_{c1}-$Cool_1$ (y_{c1}=0) after the first 50s. Apparently, this decision causes infeasibility of T_1 between the sequences, as shown in Figure 6.17, that for the remaining sequences the controller is switched back to m_{c1}-T_1, as shown in Figure 6.16.

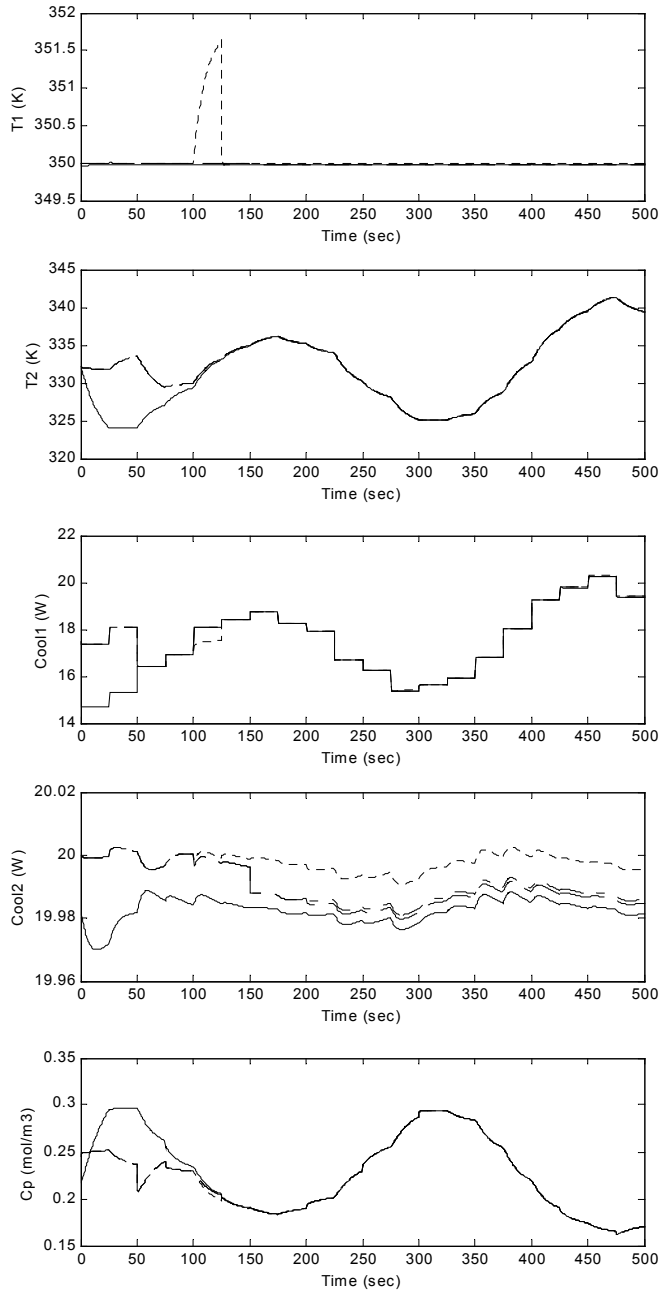


Figure 6.17 Overall Process Responses: Sinusoidal Disturbance Profile

——: opΔt=500, - - - : opΔt=50, — - : opΔt=150, ——: opΔt=300

6.6.3 Random Disturbance Profiles

In this case, the uniform random profiles of feed temperature and concentration are generated between the bounds 298K≤T_f (t) ≤315K and 19.5mol/m^3≤ C_f(t)≤ 21mol/m^3 every 25s and 50s, respectively. Hence the sampling interval is 25s, and the profiles are shown in Figure 6.18.

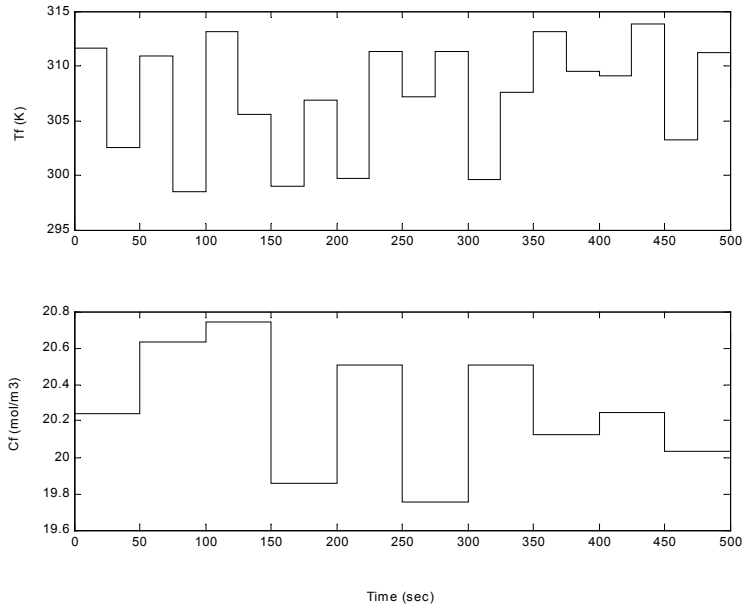


Figure 6.18 Random Disturbance Profile

The assessment results are summarised in Table 6.3. Similar to the sinusoidal case, the sequential framework finds purely parallel process structure ([y_{p1}, y_{p2}, y_{p3}, y_{p4}] = [1, 1, 0, 1]) and controller structure ([y_{c1}, y_{c2}] =[1, 0]) at all cases. The disturbance dynamics are spread evenly over the time horizon. Therefore, the disturbance time averages are [T_f C_f] = [306, 20.25], which are higher than the nominal values used by the non-sequential version. Consequently, the time average profits are lower than the non-sequential ones.

Table 6.3 Controllability Assessment Results: Random Disturbance Profile

	Non-sequential	Sequential approach			
		opΔt=50s (average)	opΔt=150s (average)	opΔt=300s (average)	opΔt=500s
Φ \$/h	106.5420	91.3940	91.3940	91.3486	92.0089
T_{fN} K	300	307.0515	307.0515	307.0515	307.0515
C_{fN} mol/m^3	20	20.2637	20.2637	20.2637	20.2637
Q_{f1} m^3 (y_{p1})	0.3561 (1)	0.3008 (1)	0.3008 (1)	0.3071 (1)	0.3027 (1)
Q_{f2} m^3 (y_{p2})	0.2882 (1)	0.2658 (1)	0.2658 (1)	0.2701 (1)	0.2670 (1)
Q_3 m^3 (y_{p3})	0.0500 (1)	0.0000 (0)	0.0000 (0)	0.0000 (0)	0.0000 (0)
Q_4 m^3 (y_{p4})	0 3061 (1)	0.3008 (1)	0.3008 (1)	0.3071 (1)	0.3027 (1)
α_1 (y_{c1})	2.0000 (1)	2.0000	2.0000	2.0000	2.0000
α_2 (y_{c2})	2.0000 (0)	2.0000	2.0000	2.0000	2.0000
V_1 m^3	4.5000	4.5000	4.5000	4.5000	4.5000
V_2 m^3	4.5000	4.5000	4.5000	4.5000	4.5000
CPU time (min)	5.0230	6.6360	6.6360	12.8297	12.7665

The above table shows similar computational performances between 50s and 150s windows, and between 300s and 500s windows. This indicates similar quality of information contained within the respective windows. In this case, the disturbance dynamics are distributed evenly, and they require a large window (300s) to yield feasible designs.

The figures 6.19-6.20 show that over the sequences, the solutions of different windows tend to converge to the solutions of 500s window. This tendency is also shown by the process responses, subject to the design sequences application, as shown in Figure 6.21. Similar to the previous sinusoidal disturbance case, smaller windows produce infeasible results prior the completion of the full optimisation window.

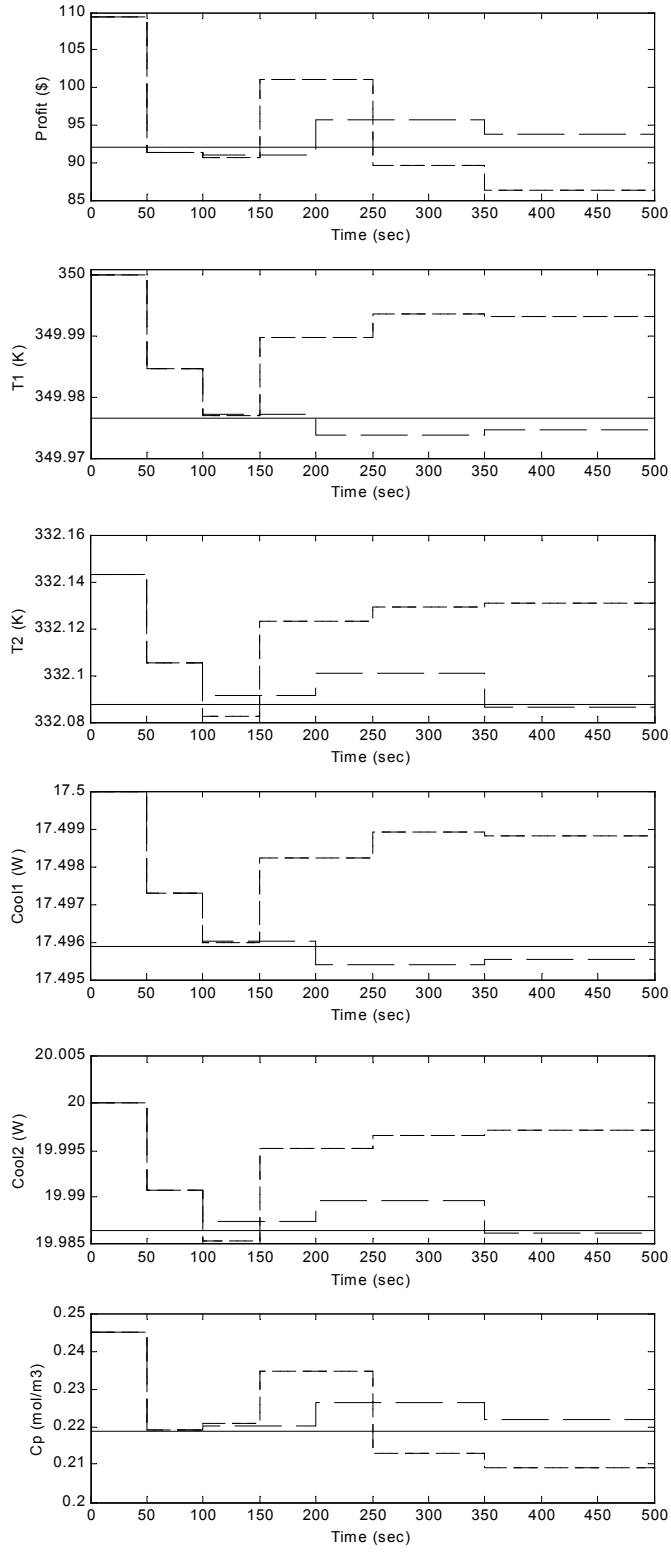


Figure 6.19 Profit and Initial Condition Sequences: Random Disturbance Profile

——— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

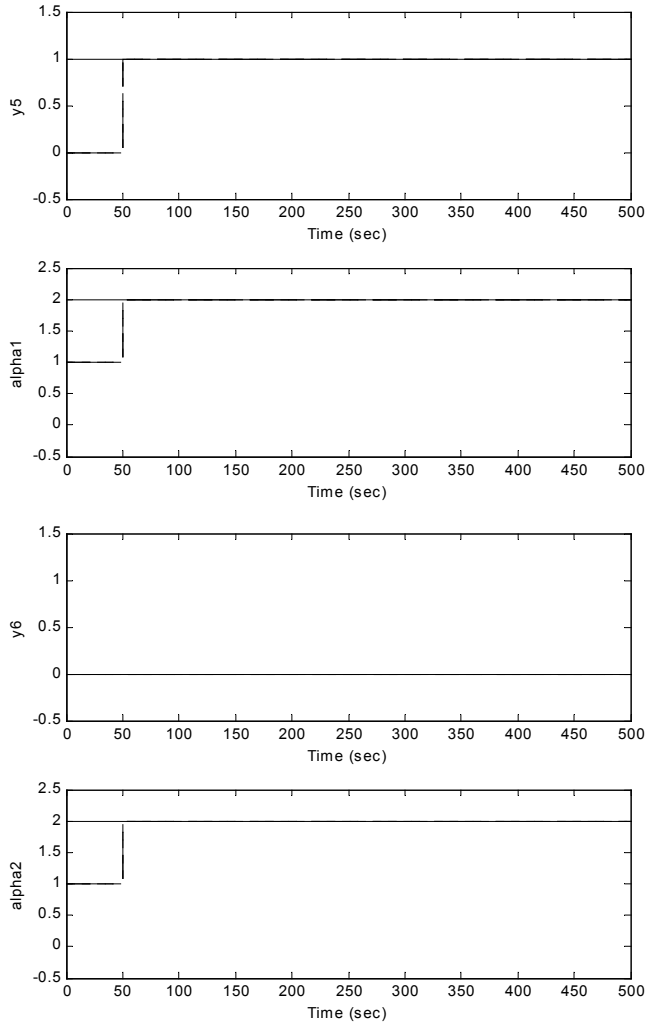


Figure 6.20 Controller Structure: Random Disturbance Profile

——— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, —— : opΔt=300

Figure 6.21 shows that the 50s and 150s windows deliver repeatedly infeasible designs subject to the disturbances, especially at $Cool_1$, $Cool_2$ and C_p. The design sequences found by the 300s window, on the other hand, keep the process feasible and close to the one produced by the 500s window, especially after the whole window is built. This case highlights the trade-off between the size of the optimisation window and the computational cost.

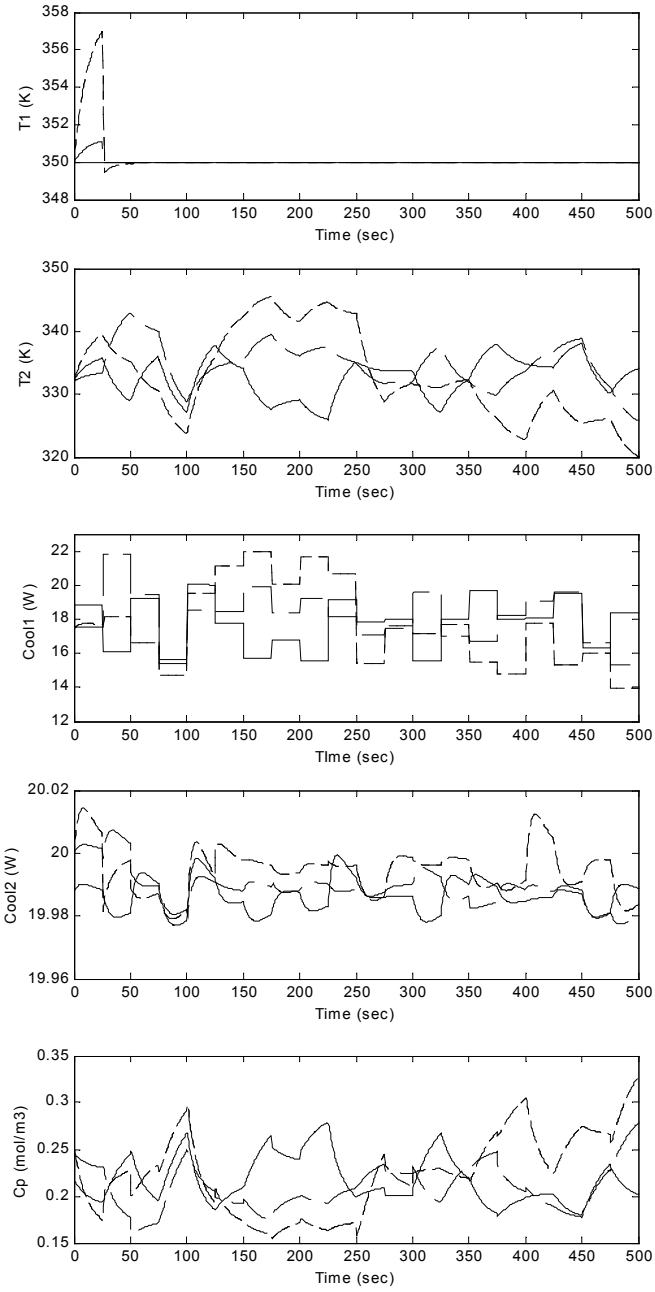


Figure 6.21 Overall Process Responses: Random Disturbance Profile

—— : opΔt=500, - - - : opΔt=50, — - : opΔt=150, ——— : opΔt=300

6.6.4 Final Discussions

These case studies demonstrated the application of the sequential framework to produce sequences of process control strategies to handle the affecting disturbance profiles over a time horizon. The changes of control parameters within the sequences were in line with the existing application of predictive and unfalsification methods as discussed in sections 6.3.4 and 6.3.5. This study has extended the problem to process and controller structures.

The use of Distributed Control System (DCS) and the existence of unit processes with inherent superstructure have been motivating the possibility of structure changes in real cases. For a given sequence, DCS provides easier switching between predefined controller structures. It should be noted that the switching is expected to occur due to major disturbance change and remain over a relatively long period; as frequent changes are easier handled by adjusting controller parameters.

There are unit processes built with inherent superstructure, to accommodate wide variation of economic objectives. In this study, process structure is the dominant factor in achieving the optimum economic objective; while the controller structure and the controller parameters fine-tune the controllability. Therefore, once selected, the process structure is less likely to change over the time horizon unless the economic policy changes. The logical application of this approach is to provide an *a-priori* map of control strategy over the time horizon of interest.

This also highlights the unfalsification priority; firstly the control parameters, then the controller structure, and lastly the process structure as well as parameters. This is consistently shown in the case study, without explicit priority enforcement within the framework. It would be interesting to see the implication of different unfalsification priorities as well as the changes of objectives in further works.

The case studies show significant differences in computational cost between the sequential framework and the non-sequential version. At this stage, the cost of solving sequential framework is still high, due to the number of piecewise dynamics within the

respective optimisation window. Accordingly, smaller windows provide faster computation. However, this risks process infeasibility between the sequences, a well-known risk of adopting the adaptive concept. Further exploration of the trade-off between the window size, computational cost and the degree of acceptable feasibility is therefore recommended for future works.

The capability of handling general disturbance profiles suggests a greater use of process operational history. In combination with the current computational cost, the framework is more suitable for off-line and long-term process planning based on the history, rather than for on-line operation. This will be illustrated further in the framework application to a real industrial unit process with inherent superstructure in Chapter 7.

The potential extension of this feature includes the representation of the process with the measurement data, instead of mathematical models. For further works, the augmented process and disturbance information (Chen *et al.*, 2000; Kam *et al.*, 2001; Mekarapiruk and Luus, 1994; Quintero and Quiennec, 2002; Sun and Tsao, 2001) may be used within the framework.

The variations of the disturbances and the control sequences may causes the discontinuity in process responses, as shown in Figure 6.13, 6.17 and 6.21. The source of these discontinuities is the nature algebraic variables in the corresponding DAE. In this study, these are the manipulated variables m_{c1} and m_{c2} and the output variables $Cool_1$ and $Cool_2$. These variables do not have derivative terms, as shown in the mass matrix arrangement of the DAE solution (see Chapter 3, equation (3.13)). Therefore, unlike the state variables, the algebraic variables do not have reset actions to smoothen the changes in their profile if any new external forces, such as disturbances, are applied. To alleviate this problem, the algebraic variables have to be changed into state variables by deriving the respective derivatives. This procedure requires major modification and thorough analysis in process model, hence recommended for future works.

As discussed earlier, the dynamic optimisation within the framework maintains its dynamic programming approach, which solves the final values, then updates the dynamic path backward to the initial values. However, the forward dynamic optimisation would provide practical tools for ongoing and long-term process planning and design. To allow smooth dynamic responses over the sequences, the optimisation problem should be reformulated. This includes constraining the initial values to the final values of the previous sequences, and shaping the dynamic profiles to achieve the optimum objective. Further consideration on the degree of process feasibility between optimisation sequence is required, as well as the effective solution of the path constraints problem (Bojkov and Luus, 1994; Vassiliadis *et al.*, 1994; Wang and Shieh, 1997). The scope of this development is very specific, and deserves further works.

The piecewise process dynamics computation does not provide the steady state data originally used by the redundancy analysis and elimination procedure, which supports the geometric computation of r-OCI. This leads to the relaxed tolerance in collinearity detection among the functional variables, which is determined by trial and error. Further study is required to resolve this problem.

6.7 Conclusion

This chapter addressed the controllability assessment of general disturbance profiles, focusing on the measurable disturbances. To enable dynamic assessment, the disturbance profiles were sampled, and the associated piecewise process responses were assessed within the Dynamic Operability Framework. This enabled controllability assessment to focus on general disturbances profiles, such as those collected from plant historical data.

To reduce the conservatism in the design, the framework algorithm has been modified to progress sequentially within an optimisation window over a time horizon. The unfalsification approach has been adopted; in which the recommended designs for the

next sequence is based on the analysis of the sampled disturbances and the associated piecewise process responses within the preceding window.

The performance of the framework has been demonstrated through case studies involving a nonlinear exothermic chemical reaction in a superstructure, affected by various disturbance profiles. The changes in process profit, controller parameters and structure due to different optimisation windows have been presented. These highlighted the trade-off between the window size, computational cost, and the feasibility of the design.

At this stage, the computational cost due to the piecewise dynamics as well as structural problem within the sequential framework is higher than the non-sequential version. Therefore, the sequential framework is suitable for off-line and long-term process planning based on process history, rather than for on-line design. This will be illustrated further in a real industrial case presented in Chapter 7.

The disturbance sampling approach leads to the idea of using process data, instead of a mathematical model within the framework. This is recommended for further study, along with the problem reformulation to adopt forward dynamic programming approach and effective solution of the path constraints problem. The trade-off between the window size, computational cost and the degree of feasibility are inherent in this problem and deserve further exploration.

6.8 Nomenclature

Acronyms		
AOS_θ	:	Achievable Output Space due to disturbances
AR	:	Auto Regressive
BB	:	Branch and Bound
CSTR	:	Continuous Stirred Tank Reactors
DAE	:	Differential Algebraic Equations
DCS	:	Distributed Control System
DMC	:	Dynamic Matrix Control
DOF	:	Dynamic Operability Framework
DOS	:	Desired Output Space
EDS	:	Expected Disturbance Space
EKF	:	Extended Kalman Filter
GPC	:	Generalised Predictive Control
r-OCI	:	Regulatory Output Controllability Index
LHS	:	Latin Hypercube Sampling
LQG	:	Linear Quadratic Gaussian
HSS	:	Hammersley Sequence Sampling
MCS	:	Monte Carlo Sampling
MINLP	:	Mixed Integer Nonlinear Programming
MPC	:	Model Predictive Control
NLP	:	Nonlinear Programming
PCA	:	Principal Component Analysis
PI	:	Proportional – Integral
PID	:	Proportional – Integral – Derivative
PVAR	:	Profit Variation
r-OCI	:	Regulatory Output Controllability Index
SQP	:	Sequential Quadratic Programming

Variables and Matrices		
$\mathbf{A_p}$, $\mathbf{B_p}$, $\mathbf{B_d}$, $\mathbf{C_p}$, $\mathbf{C_d}$, $\mathbf{D}$	:	Linearised Prediction state-space matrices
C	:	Concentration of A in CSTR (mol/m^3)
Cool	:	Amount of heat removed from CSTR by coolant
$\mathbf{E}$	:	Set of equality constraints
E_c	:	Activation energy
$\mathbf{E_\theta}$	:	Expected values of the variations due to $\boldsymbol{\theta}$
$\mathbf{g_j}$	:	Inequality constraints
$\mathbf{g_{jd}}$	:	Logic propositions
$\mathbf{h_i}$	:	Equality constraints
H_2	:	H_2 norm

Variables and Matrices		
H_∞	:	H_∞ norm
$\mathbf{I}$	:	Set of inequality constraints
$\mathbf{i}$	:	Indices of equality constraints
$\mathbf{j}$	:	Indices of inequality constraints
$\mathbf{K}$	:	Controller parameters set
k_o	:	Reaction constant
L	:	Control horizon
M	:	Measurement set
m	:	Coolant flow rate (m^3/sec)
op	:	Optimisation window
$\mathbf{P}$	:	Set of perturbed process data
$\mathbf{p}$	:	Vector of process parameters
Q	:	Liquid Flow rate (m^3/sec)
Q_ξ	:	Covariance matrix
R	:	Prediction horizon
T	:	Temperature of the reactor (°K)
t	:	Time (sec)
t_o	:	Initial time (sec)
t_{ow}	:	Optimisation sequence time (sec)
t_f	:	Final time (sec)
t_s	:	Disturbance sampling time (sec)
$\mathbf{U}$	:	Set of manipulated variable values
$\mathbf{u}$	:	Vector of manipulated variables
V	:	CSTR volume (m^3)
$\mathbf{w}$	:	Vector of output variables
$\overline{\mathbf{w}}$	:	Vector of functional measurement variables
ws	:	Optimisation sequences
$\mathbf{x}$	:	Vector of state variables
$\mathbf{y}$	:	Vector of binary decision variables
$\mathbf{Z}$	:	Set of continuous decision variable values
$\mathbf{z}$	:	Vector of continuous decision variables
$\overline{\mathbf{z}}$	:	Continuous decision variables

Greek letters		
α	:	Scaling factor of controller parameters
β	:	Performance specification
Δt	:	Optimisation sequence
Θ	:	Set of all disturbance realisations
Φ	:	Objective function
θ	:	Vector of disturbances and uncertainties
θ_e	:	Vector of measurable external disturbances
ξ	:	Zero mean white noise

Superscripts		
k	:	Critical values
l	:	Lower bounds
u	:	Upper bounds
N	:	Nominal values
*	:	Optimal solution from the last iteration

Subscripts		
ss	:	Initial conditions
1	:	First CSTR
2	:	Second CSTR
f	:	Feed
i	:	Equality index
j	:	Inequality index
jd	:	Logic/binary proposition index
e	:	Measurable disturbances
q	:	Quadrature points
s	:	Measured data

6.9 References

Acevedo, J. and Pistikopoulos, E. N. (1996). "A Parametric MINLP Algorithm for Process Synthesis Problems under Uncertainty", Industrial and Engineering Chemistry Research, 35, 147 - 158.

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Bahri, P. A., Bandoni, A. and Romagnoli, J. (1996). "Operability Assessment in Chemical Plants", Computers and Chemical Engineering, 20, S787 - S792.

Bandoni, A. J., Barton, G. W. and Romagnoli, J. A. (1994). "On Optimizing Control and the Effect of Disturbances: Calculation of the Open Loop Back-Off", Computers and Chemical Engineering, 18S(Supplement), S505 - S509.

Bansal, V., Perkins, J. D. and Pistikopoulos, E. N. (2002). "Flexibility Analysis and Design Using a Parametric Programming Framework", AIChE Journal, 48(12), 2851-2686.

Bemporad, A. (1998). "Reducing Conservativeness in Predictive Control of Constrained Systems with Disturbances", The 37th IEEE Conference on Decision & Control, Tampa, Florida, USA.

Bequette, B. W. (1991). "Nonlinear Control of Chemical Processes: A Review", Industrial and Engineering Chemistry Research, 30, 1391 - 1413.

Bernardo, F. P. and Saraiva, P. M. (1998). "Robust Optimization Framework for Process Parameter and Tolerance Design", AIChE Journal, 44(9), 2007 - 2017.

Bhatia, T. K. and Biegler, L. T. (1998). "Dynamic Optimization for Batch Design and Scheduling with Process Model Uncertainty", Industrial and Engineering Chemistry Research, 36(9), 3708 - 3717.

Bock, H. G., Diehl, M., Schloder, J. P., Algower, F., Findeisen, R. and Nagy, Z. (2000). "Real Time Optimization and Nonlinear Model Predictive Control of Processes Governed by Differential Algebraic Equations", IFAC International Symposium on Advanced Control of Chemical Processes (ADCHEM 2000), Pisa, Italy.

Bojkov, B. and Luus, R. (1994). "Time - Optimal Control by Iterative Dynamic Programming", Industrial and Engineering Chemistry Research, 33, 1486 - 1492.

Brozenec, T. F., Tsao, T. C. and Safonov, M. G. (2001). "Controller Validation", International Journal of Adaptive Control and Signal Processing, 15(5), 431 - 444.

Byrnes, C. I. and Isidori, A. (2000). "Output Regulation for Nonlinear Systems: An Overview", International Journal of Robust and Nonlinear Control, 10(5), 323-337.

Campbell, J. C. and Rawlings, J. B. (1998). "Predictive Control of Sheet- and Film-Forming Processes", AIChE Journal, 44(8), 1713-1723.

Cervantes, A. and Biegler, L. T. (1997). "Large-Scale DAE Optimization Using Simultaneous Nonlinear Programming Formulations", AIChE Journal, 44, 1038 - 1050.

Chen, X., Fukuda, T. and Young, K. D. (2000). "A New Nonlinear Robust Disturbance Observer", Systems & Control Letters, 41, 189 - 199.

Davison, D. E., Kabamba, P. T. and Meerkov, S. M. (1999). "Robustness with Respect to Disturbance Model Uncertainty - Analysis and Design", Ieee Transactions on Automatic Control, 44(8), 1556-1559.

Davison, D. E., Kabamba, P. T. and Meerkov, S. M. (2000). "Robustness with Respect to Disturbance Model Uncertainty: Theory and Application to Autopilot Performance Analysis", Mathematical Problems in Engineering, 6(2-3), 267-304.

Dimitriadis, V. D. and Pistikopoulos, E. N. (1995). "Flexibility Analysis of Dynamic Systems", Industrial and Engineering Chemistry Research, 34, 4451 - 4462.

Ekawati, E. (2002). "Systematic Controllability Assessment in Dynamic Operability Framework", The Third Inter University Postgraduate Electrical Engineering Symposium, Murdoch University, Rockingham Campus, Australia.

Ekawati, E. and Bahri, P. A. (2001). "Adaptation of Output Controllability Index within Dynamic Operability Framework", 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS 2001), Jejudo Island, Korea, Elsevier Science Ltd.

Figueroa, J. L. (2000a). "Back-Off and Model Predictive Control: A Joint Approach", IFAC International Symposium on Advanced Control of Chemical Process (ADCHEM 2000), Pisa, Italy.

Figueroa, J. L. (2000b). "Economic Performance of Variable Structure Control: A Case Study", Computers and Chemical Engineering, 24, 1821 - 1827.

Gokcek, C., Kabamba, P. T. and Meerkov, S. M. (2000). "Analysis of H-2 Performance Robustness with Respect to Disturbance Model Uncertainty", Ieee Transactions on Automatic Control, 45(8), 1541-1544.

Grossmann, I. E. and Sargent, R. W. H. (1978). "Optimal Design of Chemical Plants with Uncertain Parameters", AIChE Journal, 24(6), 1021 - 1028.

Halemane, K. P. and Grossmann, I. E. (1983). "Optimal Process Design under Uncertainty", AIChE Journal, 29, 425.

Helton, J. W. and James, M. R. (1999). "Extending H-Infinity Control to Nonlinear Systems: Control of Nonlinear Systems to Achieve Performance Objectives (Advances in Design and Control)", Society for Industrial & Applied Mathematics.

Jun, M. and Safonov, M. G. (1999). "Automatic PID Tuning: An Application of Unfalsified Control", IEEE International Symposium on Computer Aided Control System Design, Kohala Coast-Island of Hawai'i, Hawai'i, USA.

Kadali, R. and Huang, B. (2002). "Estimation of the Dynamic Matrix and Noise Model for Model Predictive Control Using Closed Loop Data", Industrial and Engineering Chemistry Research, 41, 842 - 852.

Kam, K. M., Tade, M. O., Rangaiah, G. P. and Tian, Y. C. (2001). "Strategies for Enhancing Geometric Nonlinear Control of an Industrial Evaporator System", Industrial and Engineering Chemistry Research, 40(2), 656-667.

Kim, K. J. and Diwekar, U. M. (2002). "Efficient Combinatorial Optimization under Uncertainty: 1. Algorithmic Development", Industrial and Engineering Chemistry Research, 41, 1276 - 1284.

Kosut, R. L. (2001). "Iterative Adaptive Control: Windsurfing with Confidence", Model Identification and Adaptive Control: From Windsurfing to Telecommunication, Goodwin, G., Springer-Verlag.

Krishnan, A. and Hoo, K. A. (1999). "A Multiscale Model Predictive Control Strategy", Industrial and Engineering Chemistry Research, 38, 1973 - 1986.

Kurtz, M. J. and Henson, M. A. (1997). "Input-Output Linearizing Control of Constrained Nonlinear Processes", Journal of Process Control, 7(1), 3-17.

Leyffer, S. (2001). "Integrating SQP and Branch-and-Bound for Mixed Integer Nonlinear Programming", Computational Optimization and Applications, 18, 295-309.

Liu, C. S. and Peng, H. (2000). "Disturbance Observer Based Tracking Control", Transactions of ASME, 122, 332 - 335.

Liu, M. L. and Sahinidis, N. (1996). "Optimisation in Process Planning under Uncertainty", Industrial and Engineering Chemistry Research, 35, 4154 - 4165.

Marconi, L. and Isidori, A. (2000). "Robust Global Stabilization of a Class of Uncertain Feedforward Nonlinear Systems", Systems & Control Letters, 41(4), 281-290.

Marconi, L., Isidori, A. and Serrani, A. (2002). "Input Disturbance Suppression for a Class of Feedforward Uncertain Nonlinear Systems", Systems & Control Letters, 45(3), 227 - 236.

Mekarapiruk, W. and Luus, R. (1994). "Optimal Control of Inequality State Constrained System", Industrial and Engineering Chemistry Research, 36, 1686 - 1694.

Mohideen, J., Perkins, J. D. and Pistikopoulos, E. N. (1996). "Optimal Design of Dynamic Systems under Uncertainty", AIChE Journal, 42, 2251-2272.

Nishida, N., Stephanopoulos, G. and Westerberg, A. W. (1981). "A Review of Process Synthesis", AIChE Journal, 27, 321 - 351.

Novak, Z. and Kravanja, Z. (1999). "Mixed-Integer Nonlinear Programming Problem Process Synthesis under Uncertainty by Reduced Dimensional Stochastic Optimization", Industrial and Engineering Chemistry Research, 38, 2680-2698.

Novak, Z. and Kravanja, Z. (2003). "A Strategy for MINLP Synthesis of Flexible and Operable Processes", Fourth International Conference on Foundation of Computer Aided Process Operations (FOCAPO 2003), Coral Springs, Florida, USA.

Paganini, F. (1996). "A Set-Based Approach for White Noise Modelling", IEEE Transactions of on Automatic Control, 41(10), 1453 - 1465.

Pannocchia, G. (2003). "Disturbance Modelling for Robust Control of Ill-Conditioned Processes with MPC", Journal of Process Control, 13, 693 - 701.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994a). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Dynamic and Control Structure Consideration", Industrial and Engineering Chemistry Research, 33, 1738 - 1755.

Papalexandri, K. P. and Pistikopoulos, E. N. (1994b). "Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Flexibility and Structural Controllability Aspects", Industrial and Engineering Chemistry Research, 33, 1718 - 1737.

Pistikopoulos, E. N. and Grossmann, I. E. (1988a). "Evaluation and Redesign for Improving Flexibility in Linear System with Infeasible Nominal Conditions", Computers and Chemical Engineering, 12, 841.

Pistikopoulos, E. N. and Grossmann, I. E. (1988b). "Optimal Retrofit Design for Improving Process Flexibility in Linear System", Computers and Chemical Engineering, 12, 719 - 731.

Pistikopoulos, E. N. and Grossmann, I. E. (1988c). "Stochastic Optimization of Flexibility in Retrofit Design of Linear Systems", Computers and Chemical Engineering, 12(12), 1215 - 1227.

Quintero, C. S. G. and Quiennec, I. (2002). "State and Disturbance Estimation for an Alternating Activated Sludge Process", IFAC 15th Triennial World Congress, Barcelona, Spain.

Raspanti, C. G., Bandoni, J. A. and Biegler, L. T. (2000). "New Strategies for Flexibility Analysis and Design under Uncertainty", Computers and Chemical Engineering, 24, 2193 - 2209.

Rigopoulos, A., Arkun, Y. and Kayihan, F. (1997). "Identification of Full Profile Disturbance Models for Sheet Forming Processes", AIChE Journal, 43(3), 727 - 739.

Rooney, W. C. and Biegler, L. T. (2001). "Design for Model Parameter Uncertainty Using Nonlinear Confidence Regions", AIChE Journal, 47(8), 1794-1804.

Rooney, W. C. and Biegler, L. T. (2003). "Optimal Process Design with Model Parameter Uncertainty and Process Variability", AIChE Journal, 49(2), 438 - 449.

Safonov, M. G. (1996). "Focusing on the Knowable", Control Using Logic-Based Switching, Morse, A. S., Berlin, Springer-Verlag, 224 - 233.

Safonov, M. G. and Tsao, T. C. (1997). "The Unfalsified Control Concept and Learning", IEEE Transactions of on Automatic Control, 42(6), 843 - 847.

Samsatli, N. J., Papageorgiou, L. G. and Shah, N. (1998). "Robustness Metrics for Dynamic Optimisation Models under Parameter Uncertainty", AIChE Journal, 44(9), 1993 - 2006.

Sun, Z. and Tsao, T. C. (2001). "Control of Linear Systems with Nonlinear Disturbance Dynamics", ACC, Arlington.

Swaney, R. E. and Grossman, I. E. (1985a). "An Index for Operational Flexibility in Chemical Process Design. Part I: Formulation and Theory", AIChE Journal, 31, 621 - 630.

Swaney, R. E. and Grossman, I. E. (1985b). "An Index for Operational Flexibility in Chemical Process Design. Part II: Computational Algorithm", AIChE Journal, 31, 631 - 641.

Tenny, M. J., Rawlings, J. B. and Wright, S. J. (2002). "Closed-Loop Behaviour of Nonlinear Model Predictive Control", University of Winsconsin, Madison, WI, USA.

Tjoa, I. B. and Biegler, L. T. (1991). "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential - Algebraic Equation Systems", Industrial and Engineering Chemistry Research, 30, 376 - 385.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Veres, S. M. (1999). "Iterative Identification and Control Redesign Via Unfalsified Sets of Models: A Basic Scheme", International Journal of Control, 72(10), 887 - 903.

Vinnicombe, G. (2001). "Uncertainty and Feedback : H_∞ Loop-Shaping and the ν-Gap Metric", London, Imperial College Press.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10(2), 185 - 194.

Wang, F. S. and Shieh, T. L. (1997). "Extension of Iterative Dynamic Programming to Multiobjective Optimal Control Problems", Industrial and Engineering Chemistry Research, 36, 2279 - 2286.

Woodley, B. R., How, J. P. and Kosut, R. L. (1999). "Direct Unfalsified Controller Design - Solution Via Convex Optimization", American Control Conference (ACC), San Diego, CA.

Zhou, K. and Doyle, J. (1998). "Essentials of Robust Control", Prentice Hall, Inc.

Chapter 7

Controllability Assessment of an Industrial Five-Effect Evaporator Process

7.1 Introduction

In this chapter, the Dynamic Operability Framework (DOF) is applied for systematic controllability assessment of the industrial five-effect liquor-burning evaporator associated with the Bayer process in an Alumina refinery. The process is characterised by a strong interaction between liquor levels, densities and temperatures, especially in the last effect. These dynamics are represented by an index-2 Differential Algebraic Equation (DAE) model, where the uncertain parameters are estimated based on validation against plant data. The establishment and validation of the model has been an equally major part of the work, together with the formulation of controllability problem and its assessment. The controllability problem is formulated within a superstructure, which involves the selection from several alternatives of steam supplies and vapor circulation, as well as several alternative pairs of feedback and feedforward loops. The controllability assessment covers both the worst-case and the general disturbances characterisations over 24 hours operational period. Both cases justify the

four–effect operation and simple multi-loop control strategy to achieve the optimum process economy.

The early results of this problem has been published in a conference paper titled "Controllability Analysis of Industrial Five Effect Evaporator System", presented at the Fourth International Conference on Foundations of Computer Aided Process Operations (FOCAPO 2003), Coral Springs, Florida, USA, 2003.

This chapter is organised as follows: In section 7.2, the industrial implementation of the Dynamic Operability Framework is outlined. In section 7.3, the relevant researches in general and industrial evaporator processes are briefly reviewed, concentrating on the modeling and process control aspects. It then leads to the presentation of the liquor-burning evaporator process in section 7.4. The process model is given in section 7.5. The controllability specifications are described in section 7.6. The assessment results are presented in section 7.7, and finally, the chapter is concluded in section 7.8.

7.2 The Industrial Implementation of the Dynamic Operability Framework

The previous chapters presented the development of the controllability assessment methodology within the Dynamic Operability Framework (DOF) and demonstrated the implementations on four academic models. Therefore, the next obvious stage is to implement the framework into a real industrial process.

It has been accepted that a well-designed process control system provides better performances, such as economic profit, safety and favorable dynamics. These are crucial aspects in process industries, which motivate the industry's increasing interest in control applications. However, there is a lack of clear relationships between the process economy and process control, while both may also include vast amounts of empirical considerations. It is the intention of the framework to bridge this gap, and provide a systematic justification of the optimum control strategies. The optimisation

nature of the framework would accommodate various considerations around the process, which are used to determine the optimum operational structures and parameters for both process and control system.

The industrial process considered in this study is the five-effect evaporator at the liquor-burning unit, associated with an alumina refining. The role of the process is to concentrate the spent liquor, in preparation of liquor burning to remove organic impurities. This unit process is inherently a superstructure, which accommodates various operational effect and control loop alternatives.

Given a range of targeted product densities, temperatures, as well as the other allowable operational specifications, it would be valuable to assess the available process and controller designs to achieve the best process economy and control performances. This controllability assessment considers two cases of disturbance characterisations. The first is the worst-case assumptions, where the disturbances are assumed as the step functions with magnitudes varying within the specified ranges. The second case considers the general disturbance profile, which is taken from the plant historical data over a 24 hours time horizon.

The assessment honors the confidential properties of the industry. These include the physical property of the liquor components, as well as the actual values of plant historical data, process parameter determined from the data, all process variables and the respective bounds. Accordingly, the values are presented in the relative percentage of the respective reference values.

In the following sections, the background of evaporator process is presented. It is followed by the brief review of various modeling, identification and control strategies applied in the evaporator process. Then the description gets more specific on the liquor burning evaporator unit, for which the controllability is assessed in this study.

7.3 Review of Industrial Evaporator Process Control

7.3.1 Introduction of Industrial Evaporator Process

The evaporation process originated in ancient times as people started to evaporate seawater to obtain salt with the help of solar and wind energy. Nowadays, the process of concentrating a weak solution of a nonvolatile material by means of evaporation and solvent removal is a common operation in many process industries.

These evaporators consume large amounts of steam as the heat source. The solutions are heated to reach the boiling point and start the evaporation. Since many materials cannot tolerate high temperatures, their boiling points are kept low by operating the evaporators at reduced pressure. The amount of vapor produced per unit consumption of steam, which represents the increase in concentration, can be adjusted by cascading the evaporator stages (effects). The incentive for multi-effect operation is that the vapor product of one effect can be used as the energy source for the other effects, and so forth. In all cases, the energy content of the steam is eventually lost to the environment (Liptak, 1999).

The products can be a concentrated solution, such as fruit juices; or a clean and condensed solvent, as in the seawater desalination process (Liptak, 1999). It may also be an intermediate product, such as the spent liquor in alumina processing through the Bayer process.

7.3.2 Review of Modelling, Identification and Control of Evaporator Process

In an effort to increase the benefit of this energy intensive process, a significant number of studies on dynamic process modeling, identification and control of industrial evaporators have been performed over the last two decades. The relevant

studies are summarised, *inter alia*, in Table 7.1. These studies deal with highly nonlinear and ill-conditioned dynamic models, high interaction among process variables, significant uncertainties of process parameters, as well as limited online measurements.

The analytical nonlinear evaporator models are typically derived based on mass and energy balances, embedded with the thermodynamic properties of liquid and vapor (Cardoso and Dourado, 1998; Chenery, 1997; Kam and Tade, 1998; 2000; Kam *et al.*, 1998; Khan *et al.*, 1998; Kundergi and Nataraj, 1994; Newell and Lee, 1989; To *et al.*, 1995; To *et al.*, 1998b; Wang and Cameron, 1994; Winchester and Marsh, 1999). The treatments of these complications include model reduction and linearisation (Cardoso and Dourado, 1998; Montano *et al.*, 1991; Winchester and Marsh, 1999).

Table 7.1 Applied Control Strategies on Evaporator Processes

Authors	Control strategies				
	Method	Feed-back	Feedfor ward	Multi-loops	Multi-variables
Newell & Lee (1989)	GMC	√	√	√	√
Montano *et al.* (1991)	I/OLC	√	√	√	
Wang & Cameron (1994)	2 step GMC		√	√	
Kundergi & Nataraj (1994)	QFT	√			√
Elhaq *et al.* (1999)	GPC	√	√	√	
Nielsen *et al.* (1996)	Multi-loops PI	√		√	
Chenery (1997)	Multi-loops PI			√	
Cardoso & Dourado (1998)	μ-synthesis	√			√
Winchester & Marsh (1999)	Multi-loops PI	√		√	
To *et al.* (1995; 1998b)	I/OLC GLC	√		√	
Kam *et al.* (1998; 2000; 1998)	I/OLC GLC, IOIMC	√		√	

Over time, the capability to solve the models has been increasing with advances in numerical computing, which is demonstrated by the use of symbolic programming (Kam and Tade, 1998; 2000; Kam *et al.*, 1998; To *et al.*, 1995; To *et al.*, 1998b) and distributed parameter models (Stefanov and Hoo, 2003).

In the process identification area, Young and Allen (1994; 1995a; 1995b) reported that the distributed-parameter models gave no improvement over the simpler lumped-parameter counterparts. Later, Elhaq *et al.* (1999) presented the agreement between the results of the applied Controlled Auto Regressive Integrated Moving Average (CARIMA) identification method with the lumped analytical model.

Various control methods have been associated with these models, as summarised in Table 7.2. These include the Input-Output linearisation and the variants (Kam and Tade, 1998; 2000; Kam *et al.*, 1998; Montano *et al.*, 1991; To *et al.*, 1995; To *et al.*, 1998b), the Generic Model Control (GMC) (Newell and Lee, 1989; Wang and Cameron, 1994), the Generalised Predictive Control (GPC) (Elhaq *et al.*, 1999), Internal Model Control (IMC) (Kam *et al.*, 2001), adaptive neural network (Mills *et al.*, 1999) and fuzzy (Newell and Lee, 1989) approaches.

Most of the above control strategies are applied on the fixed process and controller structure. The discussion about controller structure selections, only reported in a small number of references, which include the heuristic selection of process and controller structure (Nielsen *et al.*, 1996), the RGA analysis (Winchester and Marsh, 1999), and the Relative Order Matrix (Kam and Tade, 2000; Wang and Cameron, 1994). These selections are performed based on the interaction between the manipulated variables and the output variables only, separate from economic considerations. The improvement of process economy due to control implementations is shown in a fixed process structure by Elhaq *et al.* (1999), and the discussion on simultaneous controllability and profitability analysis is only provided by Chenery (1997). This short list indicates the lack of the systematic and simultaneous assessment of process controllability and profitability involving several alternative process and controller structures, which therefore deserves further study.

Table 7.2 Process and Controller Structures of Evaporation Process

Authors	Structure	Variables			
		θ	**x**	**u**	**w**
Newell and Lee (1989), Wang & Cameron (1994), Chenery (1997)	Three-effects, nonlinear model		h_P C_P P_P	Q_P P_S M_{HF} M_{CW}	h_P C_P
Montano *et al.* (1991)	Double-effects, linearised model	Q_F C_F M_S	M_P C_P E_P	Q_F.	C_P
Kundergi & Nataraj (1994)	Three-effects, linearised model			P_S Q_P	C_P h_P
Khan *et al.* (1998)	Three-effects, nonlinear model	Q_F C_F T_F	M_P C_P		
Elhaq *et al.* (1999)	Five-effects, nonlinear model	Q_F Q_P		M_S M_V	C_P M_V
Nielsen *et al.* (1996)	Five-effects, nonlinear model	M_S Q_F C_F	M_S M_V M_F	M_S Q_P M_V	C_P M_V
Cardoso & Dourado (1998)	Six-effects, reduced linear model	C_F T_F T_S	M_P C_P	M_F M_S	M_P C_P
Winchester & Marsh (1999)	Single-effect, nonlinear model	T_F C_F	T_P C_P	Q_P M_{CW}	T_P C_P
To *et al.* (1995; 1998b)	Three-effects, nonlinear model		h_P ρ_P T_P	Q_F M_{CW} M_{HF}	h_P ρ_P T_P
Kam and Tade (1998; 2000) , Kam *et al.* (1998)	Five-effects, nonlinear model	Q_F C_F T_F	h_P ρ_P T_P	Q_P M_V M_S	h_P ρ_P T_P

Note:
In the above table, Q are the volume flow-rates, M are the mass flow-rates, h are the liquor levels, T are the temperatures, E are the enthalpies, P are the gas pressures, C are the solid compositions, and ρ are the densities.

The subscripts F, P, S, V and HF represent the feed stream, products stream, live steam, vapor and recycle streams, respectively.

7.3.3 Comments on the Index of Evaporator Dynamic Models

Table 7.2 shows that the significant output dynamics in the evaporator process are associated with the level inventories, densities and temperatures, or their variants, such as compositions, enthalpies, and vapor pressure. In general, these dynamics are represented in high index Differential Algebraic Equations (DAE) (Wang and Cameron, 1994).

By definition, the DAE index is equal to the minimum differentiation required to transform the original DAE into an explicit Ordinary Differential Equations (ODE) (Fabian *et al.*, 2001). In other words, a DAE contains dynamic variables that are defined by their algebraic relationships to other variables, and not by their derivatives. In simple control problems, the manipulated variables typically become the algebraic variables within the DAE, such that their dynamic profiles are generated simultaneously with the state dynamic profiles.

These algebraic variables require proper initialisation. The algebraic variables of an index-1 DAE can be initialised independently from other variables. However, as the index gets higher, determining these values become more difficult, because they are implicitly constrained within the models equations. The initial values of an index-2 DAE are constrained algebraically; while at the higher index DAE, the higher derivatives of the initial values are constrained.

The existence of the initialisation problems in a steady state evaporator calculation is well known and numerical and modeling software such as GAMS, SPEEDUP and MATLAB are available to solve them. In contrast, most dynamic solvers are restricted to ODE and index-1 DAE, which severely holds back the dynamic analyses of evaporation process.

Furthermore, the dynamic variables in high-index DAE are likely to be nonlinearly interactive, hence not control-affine (Kam and Tade, 2000). However, most of the nonlinear feedback control methods are based on control-affine assumption (Byrnes and Isidori, 2000), and therefore, are not directly applicable to these DAE models.

Due to these difficulties, the high index DAE problems in many cases are avoided altogether, or transformed into an index-1 DAE problem. One example of this is the assumption that the mass balance has been handled in a separate control strategy (Montano *et al.*, 1991; Nielsen *et al.*, 1996), which reduces the problem into one of controlling temperatures and densities only. The other examples are fixing the values of one or more dynamic variables, or assuming linear relationships among a subset of variables. This way, the model is transformed into an index-1 and control-affine DAE (Kam and Tade, 2000; Kam *et al.*, 1998; Kam *et al.*, 2001; To *et al.*, 1998a; To *et al.*, 1995; To *et al.*, 1998b; Wang and Cameron, 1994).

One alternative solution to this problem is offered by the simultaneous dynamic optimisation method. In this method, the dynamic profiles are discretised at a number of points (e.g. using finite element or orthogonal collocation methods), and then the variable values at each point are included as the optimisation solution. Therefore, the dynamic optimisation is converted into a large scale steady state Nonlinear Programming (NLP) problem (Tjoa and Biegler, 1991; Vassiliadis *et al.*, 1994). This approach is not constrained by the DAE indices anymore, but by the NLP size. Therefore, it is unsuitable for industrial applications. These cases show that solving high index DAE problems in a direct manner is an open challenge in control design and controllability analysis.

7.3.4 On Controllability Assessment of an Industrial Evaporator Process

It has been established that the controllability performances depend on the process and controller structure. In evaporator processes, this involves the trade-off between the gradual energy transfer of multiple-effects operation with the operational costs, as well as the rejection of the ill-matching manipulated and output variables. These lead to an inherently uncontrollable process and unnecessarily costly controller designs.

To address these issues, the Dynamic Operability Framework is employed for the systematic controllability assessment of evaporator process. The simultaneous

assessment of the process economy, and process structures has been demonstrated in academic models in the previous chapters. In this chapter, the method will be applied to a multiple-effect industrial evaporator, which is inherently a superstructure.

The new issue in this chapter is the high index DAE. As discussed in chapter 3, the Dynamic Operability Framework is potentially capable of solving the high index problem. The requirement of consistent initial conditions is inherently fulfilled within the framework structure, especially at the second and the subsequent outer-levels, which solve both steady state and dynamic models. With careful modeling, the steady state calculation would act as the constraints for the initial conditions required by the subsequent high index DAE. Both models should be mostly consistent, while acknowledging the conditions requiring different treatments between the steady state and dynamic models, such as control structure assignments.

Since the process model would not be control affine, this chapter is not focussing on controller design, but on selecting the suitable matching between the manipulated variables and output variables in a stabilising multi-loop control scheme. The industrial evaporator under study is a five-effect evaporator within the liquor- burning unit associated with Bayer process. The process description is given in the following section.

7.4 Liquor-Burning Unit in Bayer Process

The role of the liquor burning evaporator within the Bayer process is illustrated in Figure 7.1. The Bayer process commences by mixing the finely crushed bauxite with sodium hydroxide, then digesting them at high temperature and pressure. The result is slurry containing a mixture of dissolved aluminum oxides and bauxite residues. The slurry is cooled and the residue is discarded. The remaining supersaturated liquor is precipitated to produce the alumina trihydrate crystals, then classified according to size.

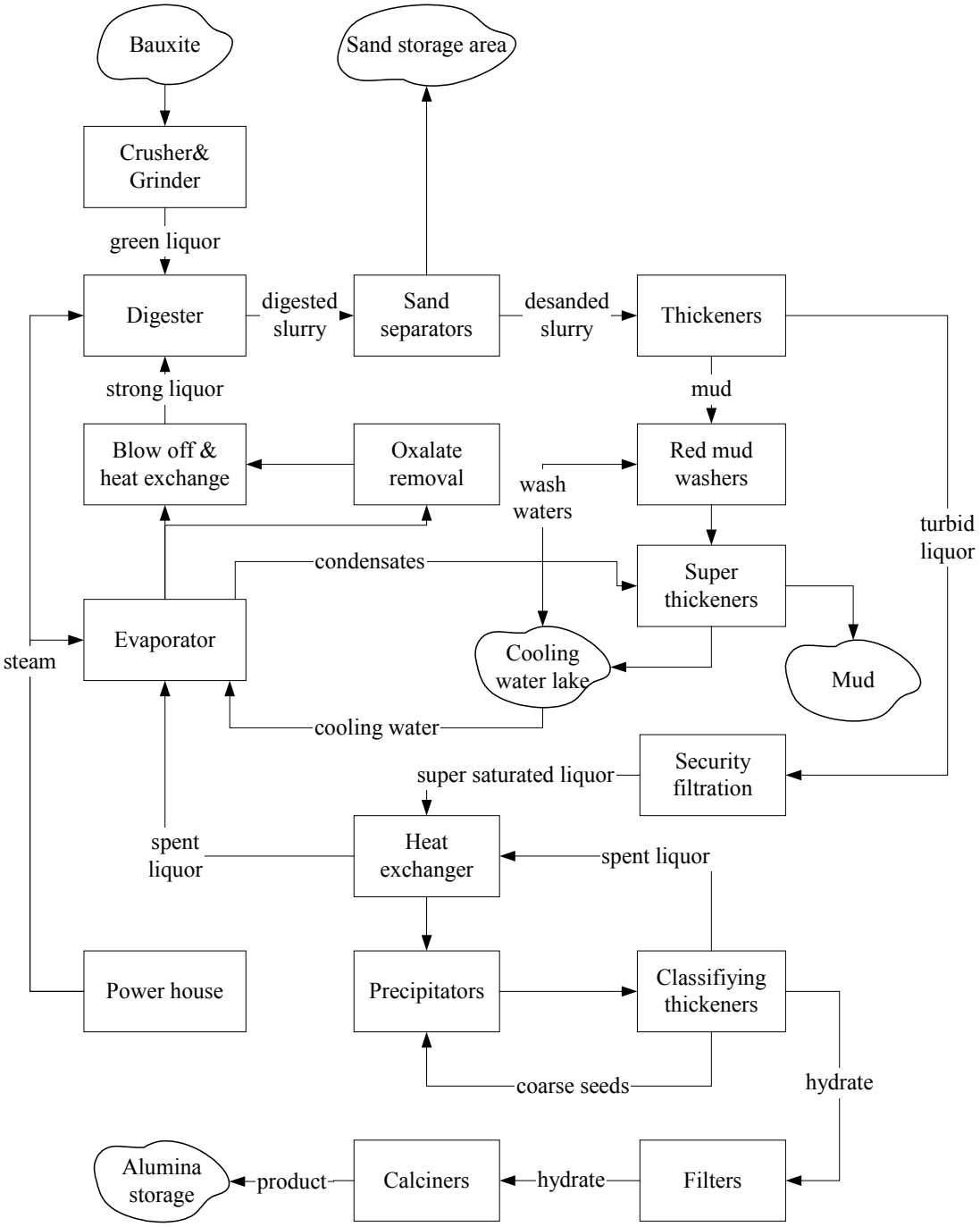


Figure 7.1 Flow Sheet of the Bayer Process (EPA, 1994; Sidrak, 2001)

The large crystals are calcined to remove the water of hydrates and then stored. The medium-sized crystals are recycled to the precipitation stage. The by-product of precipitation stage is the spent liquor, which is sent to the liquor burning evaporator. Here, the liquor is evaporated at high recycle rates to achieve a product of a specified density. Subsequently, the liquor is burned to remove the organic impurities, then recycled to the digestion stage (EPA, 1994; Sidrak, 2001).

The above description shows that the spent liquor through the evaporator is a side stream and its economic value in isolation is yet to be determined. Nevertheless, the energy consumption and the maintenance costs are significant (DISR, 2000). The evaporator is also subject to the fluctuations of liquor flow-rate, temperature and composition. The product qualities, in terms of flow-rate, density, and temperature of the concentrated liquor, determine the efficiency of the burning process. For that reason, the product temperature and composition must be maintained above the target values, and liquor inventories and surge volume must be kept within high and low limits (Sidrak, 2001).

Despite the variety of evaporator cases, the control studies of the liquor-burning evaporator in Bayer process are only reported by To, Kam and co-workers (1998; 2000; Kam *et al.*, 1998; 2001; 1998a; To *et al.*, 1995; 1998b). The corresponding controller structures are selected using a Relative Order Matrix (Daoutidis and Kravaris, 1992), without further consideration of the process economy.

To fill this gap, the Dynamic Operability Framework (DOF) is applied for systematic controllability assessment of the five-effect liquor burning evaporator. The process itself is considered as a superstructure, which contains several alternative operational effects, feedback and feedforward control loops. The objectives are to determine the structures and parameters of both process and controller, to maximise the process economy, in terms of the amount of vaporisation minus the cost of heating and condensing energy, while maintaining feasible operation. The process is represented by an index-2 Differential Algebraic Equations (DAE) model, which is described in the next section.

7.5 Five-Effect Evaporator Model

7.5.1 Superstructure Description

The liquor burning evaporator, with counter-current configuration shown in Figure 7.2, consists of one falling-film unit at stage 1, three counter-current forced circulation units at stages 2-4, and a super-concentrator at stage 5. Each stage involves the evaporation in the flash tank, mixing of feed and recycled streams, mixing of vapor, and heat transfer in the heaters at stages 1-2 and heat exchangers at stages 3-5. The live steam can be distributed to any of the heat exchangers, but at least one of the stages 3 or 4 should operate. When they are both operating, the steam supply to stage 3 is fixed, while supply to stage 4 can be manipulated for control purposes. The vapor off both stages are mixed, and then sent to the heater in stage 2. The condensation from stages 3-5 are also mixed and sent to heater 2. The vapor and condensation off stage 2 are sent to stage 1, and lastly off stage 1 to the condenser. This arrangement facilitates the alternatives of three, four or five-effect evaporator processes.

The liquor, vapor and steam flows within this counter-current superstructure are addressed by their volume flow-rate $\mathbf{Q_{Xi}}$, or mass flow-rate $\mathbf{M_{Xi}}$, where **X** is the stream and **i** is the stage number. The processes at stages 1-4 are initiated by the mixing between the feed Q_F, and the products from preceding stages $Q_{P1\text{-}3}$, with the recycle streams $Q_{I1\text{-}4}$. The mixtures are heated in heaters 1-4, and then spontaneously evaporated (flashed) in flash tanks 1-4. The concentrated liquors are divided into two groups. The first group consists of the products $Q_{P1\text{-}3}$, which proceed to the next stage. The second group consists of the recycle streams $Q_{I1\text{-}4}$, which mix with the incoming feed streams and carry on to the flashing process. The heating energy is coming from the opposite direction of the liquor flows, involving the live steam $M_{S3\text{-}5}$, vapor $M_{V3\text{-}2}$, and condensats $M_{C2\text{-}5}$.

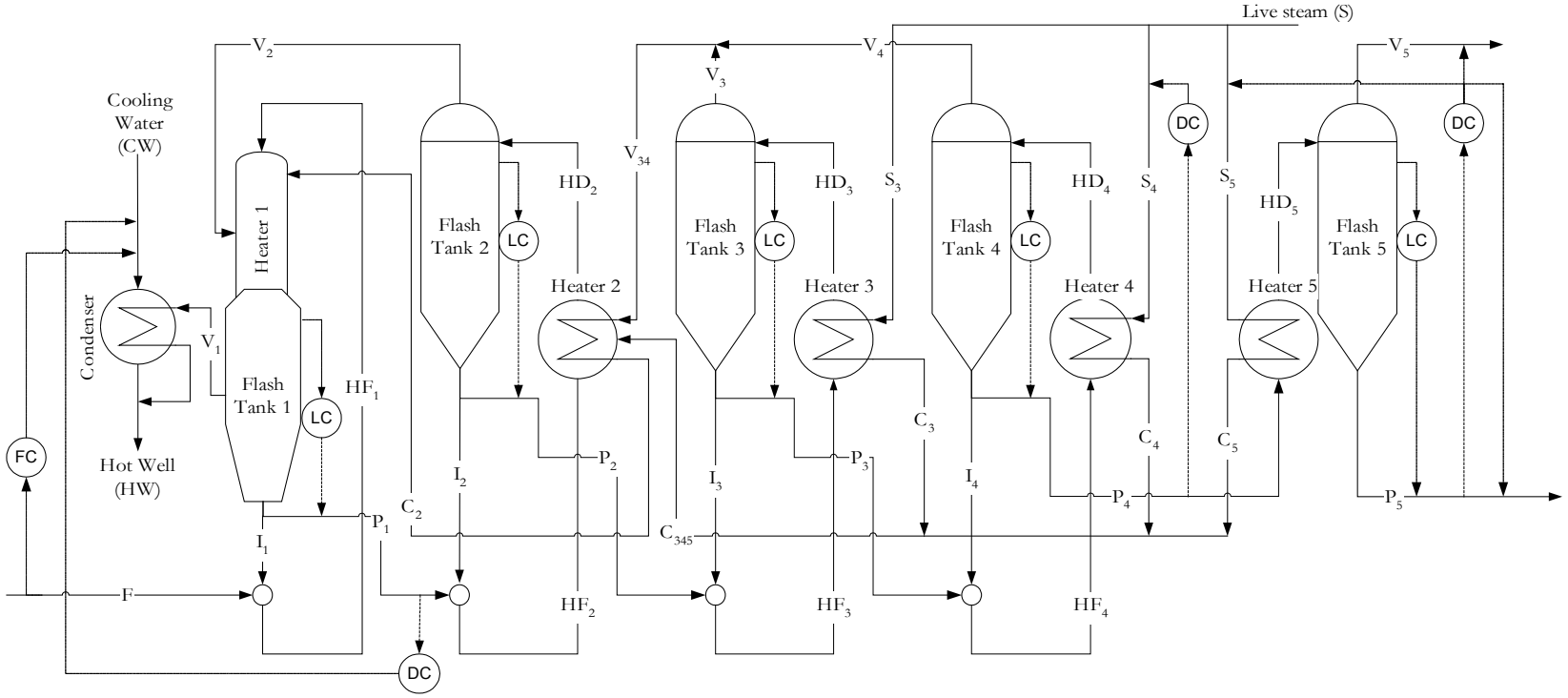


Figure 7.2 Process Flow Sheet

The measured disturbances are the feed flow-rate Q_F, feed temperature T_F, and cooling water temperature T_{CW}. The feed density ρ_F is assumed as the function of the temperature T_F and the solid composition array C_F. This array contains the solid composition of six proprietary components within the liquor. Based on the material balance, the solid compositions within the liquor around the flow sheet, C_L, are assumed to vary linearly with C_F, hence $C_L = x_L C_F$, where x_L is the solid composition ratio, and y_F=1.

The assumed state variables capturing the regulatory dynamics subject to disturbances are the liquor levels h_{P1-5}, composition ratio $y_{P1-5,}$ and enthalpies E_{P1-5} in all flash tanks. However, only the liquor levels and temperatures of all stages, h_{P1-5}, T_{P1-5}, and densities of stages 4 and 5, ρ_{P4-5} can be measured. The candidates of manipulated variables are the product flow-rate Q_{P1-5}, the live steam M_{S3-5}, the vapor M_{V5}, and the cooling water M_{CW}. The alternative matching between these measurements and manipulated variables define the control structure problems. The corresponding feedback and feedforward control-loops are defined in section 7.6.3 and illustrated in Figure 7.2.

The controllability assessment is based on process dynamics over a 24-hours time horizon. The selection of the process and controller structures is facilitated by binary variables y_p and y_C respectively. The assignments of these variables into process model and constraints are discussed further in section 7.6.2 and 7.6.3.

The independent uncertain parameters, which are the feed solid composition C_F, and the heat transfer coefficients of heat exchangers and heaters UA_{1-5}, are estimated during model validation against plant data in section 7.5.3. The effect of UA_{1-5} are considered slow relative to the assessed time horizon. Therefore, the values are assumed constant during the process dynamics. In contrast, the effect of C_F propagates throughout the dynamics of all effects. Therefore, an assumed range of this uncertainty is included in the worst-case analysis of the controllability assessment.

7.5.2 Superstructure Model

The process model in this study is based on the steady state model derived from the mass and energy balances around the flow sheet (Nooraii, 1996). The model is extended to cover the process dynamics, and to accommodate the process and controller structure selection. Both steady state and dynamic models are used within the framework. The assumptions in these models are as follows:

1. The state variables are the liquor levels h_{Pi}, solid composition ratios y_{Pi}, and temperatures T_{Pi} at each flash tanks.
2. There is no chemical reaction among solid components and between solids and solvent. Therefore, although C_F consists of six different components, their compositions increases at a specific point within the process are represented by solid composition ratio x_L. Here, L represents the subscripts F, P, and HF, which are feed, products, and heater feed streams, respectively.
3. The solution is perfectly mixed.
4. The steam enthalpy is constant and there is no steam retention.
5. Only water is evaporated.
6. The water vapor and steam are in saturated equilibrium.
7. The process is adiabatic.

The dynamic model is represented as a set of DAE. This set contains several modules, associated with the significant mass and energy balances within the superstructure. These are mixers, vapor mixers, heat exchangers, heaters, flash tanks, and condenser. The steady state version of this model is obtained by replacing the derivative terms with zero.

Mixers: i = 1, 2, 3, 4 P_0=F

$$M_{HF_i} = Q_{P_{i-1}} \rho_{P_{i-1}} + Q_{I_i} \rho_{P_i} \tag{7.1}$$

$$E_{HF_i} = \left(Q_{P_{i-1}} \rho_{P_{i-1}} E_{P_{i-1}} + Q_{I_i} \rho_{P_i} E_{P_i}\right) / M_{HF_i} \tag{7.2}$$

$$Q_{HF_i} = M_{HF_i} / \rho_{HF_i} \tag{7.3}$$

$$x_{HF_i} = \left(Q_{P_{i-1}} x_{P_{i-1}} + Q_{I_i} x_{P_i}\right) / Q_{HF_i} \tag{7.4}$$

$$C_{HF_i} = x_{HF_i} C_F \tag{7.5}$$

$$T_{HF_i} = T\left(E_{HF_i}, C_{HF_i}\right) \tag{7.6}$$

$$\rho_{HF_i} = \rho\left(T_{HF_i}, C_{HF_i}\right) \tag{7.7}$$

$$cp_{HF_i} = cp\left(T_{HF_i}, C_{HF_i}\right) \tag{7.8}$$

Heat exchangers: i = 3, 4, 5

$$M_{S_i}\left(H_{S_i} - E_{C_i}\right) = M_{HF_i}\left(E_{HD_i} - E_{HF_i}\right) \tag{7.9}$$

$$T_{HD_i} = T_{S_i} - \left(T_{S_i} - T_{HF_i}\right)\left(1 - \exp\left(\frac{-UA_i}{M_{HF_i} cp_{HF_i}}\right)\right) \tag{7.10}$$

$$E_{HD_i} = E\left(T_{HD_i}, C_{HD_i}\right) \tag{7.11}$$

$$E_{C_i} = E_C\left(T_{HD_i}\right) \tag{7.12}$$

Heaters: i = 3, 4, 5

$$M_{V_{34}}\left(H_{V_{34}} - E_{C_2}\right) + \sum_{i=3}^{5} M_{C_i}\left(E_{Ci} - E_{C_2}\right) = M_{HF_2}\left(E_{HD_2} - E_{HF_2}\right) \tag{7.13}$$

$$T_{HD_2} = T_{V_{34}} - \left(T_{V_{34}} - T_{HF_2}\right)\left(1 - \exp\left(\frac{-UA_2}{M_{HF_1} cp_{HF_1}}\right)\right) \tag{7.14}$$

$$M_{V_2}\left(H_{V_2} - E_{C_1}\right) + M_{C_2}\left(E_{C_2} - E_{C_1}\right) = M_{HF_1}\left(E_{HD_1} - E_{HF_1}\right) \tag{7.15}$$

$$T_{HD_1} = T_{V_2} - \left(T_{V_2} - T_{HF_1}\right)\left(1 - \exp\left(\frac{-UA_1}{M_{HF_1} cp_{HF_1}}\right)\right) \tag{7.16}$$

Condenser: i = 1

$$T_{HW} = T_{V_1} - \left(T_{V_1} - T_{CW}\right)\left(1 - \exp\left(\frac{-UA_C}{M_{CW} cp_{CW}}\right)\right) \tag{7.17}$$

$$M_{V_1}\left(H_{V_1} - E_C\left(T_{V_1}\right)\right) = M_{CW}\left(E_{HW} - E_{CW}\right) \tag{7.18}$$

Vapor mixers

$$M_{V_{34}} = M_{V_3} + M_{V_4} \tag{7.19}$$

$$BPE_{P_i} = BPE\left(T_{P_i}, C_{P_i}\right) \tag{7.20}$$

$$T_{V_i} = T_{P_i} - BPE_i - T_a \tag{7.21}$$

$$H_{V_i} = H\left(T_{P_i}, BPE_i\right) \tag{7.22}$$

$$H_{V_{34}} = \left(M_{V_3} H_{V_3} + M_{V_4} H_{V_4}\right) / M_{V_{34}} \tag{7.23}$$

$$T_{V_{34}} = T_{V_{34}}\left(H_{V_{34}}, BPE_{34}\right) \tag{7.24}$$

$$BPE_{34} = \left(BPE_3 + BPE_4\right)/2 \tag{7.25}$$

$$M_{V_3} = M_{V_{34}}\left(H_{V_4} - H_{V_{34}}\right) / \left(H_{V_4} - H_{V_3}\right) \tag{7.26}$$

$$M_{V_4} = M_{V_{34}}\left(H_{V_3} - H_{V_{34}}\right) / \left(H_{V_3} - H_{V_4}\right) \tag{7.27}$$

Flash tanks: i = 1, 2, 3, 4, 5

$$\frac{d\left(h_{P_i}\right)}{dt} = \left(M_{HF_i} - \left(Q_{I_i} + Q_{P_i}\right)\rho_{P_i} - M_{V_i}\right) / A_i \rho_{P_i} \tag{7.28}$$

$$\frac{d\left(x_{Pi}\right)}{dt} = \left(Q_{HF_i} x_{HF_i} - \left(Q_{I_i} + Q_{P_i}\right) x_{P_i}\right) / A_i h_{P_i} \tag{7.29}$$

$$\frac{d\left(E_{P_i}\right)}{dt} = \begin{pmatrix} Q_{P_{i-1}} \rho_{P_{i-1}} \left(E_{P_{i-1}} - E_{P_i}\right) \ldots \\ - M_{V_i}\left(H_{V_i} - E_{P_i}\right) + M_{S_i}\left(H_{S_i} - E_{C_i}\right) \end{pmatrix} / A_i \rho_{P_i} h_{P_i} \tag{7.30}$$

$$T_{P_i} = T\left(E_{P_i}, C_{P_i}\right) \tag{7.31}$$

$$\rho_{P_i} = \rho\left(T_{P_i}, C_{P_i}\right) \tag{7.32}$$

$$\Phi = \sum_{i=1}^{5} M_{V_i} - \begin{pmatrix} 4\times10^{-4}\begin{pmatrix} 0.5 * M_{S_5}\left(H_{S_5} - E_{C_5}\right)+\cdots \\ M_{S_4}\left(H_{S_4} - E_{C_4}\right)+\ldots \\ M_{S_3}\left(H_{S_3} - E_{C_3}\right) \end{pmatrix}+\ldots \\ 1\times10^{-4}\times M_{CW}\left(E_{HW} - E_{CW}\right) \end{pmatrix} \tag{7.33}$$

The nomenclature for this model is given in section 7.9. The detailed thermodynamic calculations in (7.7)-(7.8), (7.20)-(7.22), (7.32)-(7.53) are strongly nonlinear functions of the corresponding temperatures and the components of C_F. As mentioned earlier in this chapter, these functions are confidential properties of the industry.

The equations (7.31)-(7.32) provide the output variables ρ_{Pi} and T_{Pi}, as nonlinear functions of the state variables x_{Pi} and E_{Pi}. This indicates an index-1 DAE problem, since the values of ρ_{Pi} and T_{Pi} are easily updated once x_{Pi} and E_{Pi} are known. To provide efficient solutions within the framework, ρ_{Pi} and T_{Pi} are assigned as additional decision variables.

In stages 1-4, the incoming liquor feeds Q_{Pi-1} are mixed with the recycle streams of the respective stages Q_{Ii}. The exception is the last stage, where the product of stage 4 is heated and concentrated in a one-pass fashion, without any mixing. The mixture becomes the heater feed Q_{HF1}. The mixing calculations have a unique solution if, and only if, ρ_{HFi} are initialised consistently based on E_{HFi} and x_{HFi}. Here, ρ_{HFi} is tightly constrained by (7.3) and (7.7), which gives rise to an index-2 DAE problem. To solve this problem, ρ_{HFi} are assigned as the extra decision variables in the outer level.

The live steam is supplied into the heat exchangers in stages 3-5. The energy balance within these equipments can be represented by the set of equations involving the Log Mean Temperature Difference (LMTD) given in equations (7.34) and (7.35).

$$M_{S_i}\left(H_{S_i} - E_{C_i}\right) = M_{HF_i}\left(E_{HD_i} - E_{HF_i}\right) \tag{7.34}$$

$$M_{HF_i}\left(E_{HD_i} - E_{HF_i}\right) = UA_i \frac{\left(T_{HD_i} - T_{HF_i}\right)}{\ln\left(T_{S_i} - T_{HF_i}\right) - \ln\left(T_{S_i} - T_{HD_i}\right)} \tag{7.35}$$

The above equations represent the ideal counter-current heat exchange at the steady state condition. However, they do not provide a meaningful answer when $(Tsi-T_{HFi})=(Tsi-T_{HDi})$, $T_{Si} \leq T_{HFi}$, or $T_{Si} \leq T_{HDi}$, which are probable cases during a dynamic simulation (Glemmestad, 1997). The equations (7.34)-(7.35) also require unacceptably long iterations to solve in terms of T_{HDi}, hence their assignment as decision variables is not a viable option. Alternatively, the left-hand side of (7.35) is approximated by assuming negligible change in the heat capacity cp_{HFi} as follows:

$$M_{HF_i}\left(E_{HD_i} - E_{HF_i}\right) = M_{HF_i} cp_{HF_i}\left(T_{HD_i} - T_{HF_i}\right) \tag{7.36}$$

The combination of (7.35)-(7.36) provides the calculation of the outlet temperature $T_{HD3\text{-}4}$ (7.12). Using similar approach, the outlet temperatures of the heaters $T_{HD1\text{-}2}$ in (7.14) and (7.16), and condenser T_{CW} in (7.17) can be obtained.

When both stages 3 and 4 are operating, the vapor coming off these stages are mixed. The vapor mixing equations (7.19)-(7.27), coupled with the heat transfers (7.13)-(7.14) require accurate initialisation for M_{V3} and M_{V4}, as well as the vaporisation at the other stages. Therefore, M_{Vi} give raise to an index-2 DAE problem, and assigned as extra decision variables in the outer level.

The process economy $\mathbf{\Phi}$ is defined as the total amount of vaporisation minus the cost of heating and condensing energy (7.33). The function is included in the model to facilitate the profile analysis over time, subject to disturbances. The initialisation is not critical, therefore, it only gives rise to an index-1 DAE problem. Since this function is required for an objective function, $\mathbf{\Phi}$ is also assigned as a decision variable.

The state equations (7.28)-(7.30) outline the strong interactions among state variables, which propagates through the measured levels, densities and temperatures. These interactions are the major cause of the high computational cost of the assessment, which is discussed further in section 7.7. For the steady state model, all derivative terms in (7.28)-(7.30) are nullified. Within the outer level of the framework, the equations (7.7), (7.9), (7.13), (7.15), (7.18), (7.26)-(7.33) are considered as the steady state constraints, which determine the initial conditions for the corresponding DAE.

7.5.3 Model Validation

In total, the above evaporator model involves 75 variables, which are divided into 18 groups as summarised in Table 7.3.

Table 7.3 Process Variables

Variables						
Design	State	Output	Manipulated	Extras	Disturbances	Uncertainties
$Q_{I1\text{-}4}{}^{c}$	$h_{P1\text{-}5}{}^{a}$ $y_{P1\text{-}5}{}^{b}$ $E_{P1\text{-}5}{}^{b}$	$\rho_{P1\text{-}3}{}^{b}$ $\rho_{P4\text{-}5}{}^{a}$ $T_{P1\text{-}5}{}^{a}$	$Q_{P1\text{-}5}{}^{a}$ $M_{S3\text{-}5}{}^{a}$ $M_{CW}{}^{a}$	$M_{V1\text{-}5}{}^{c}$ $\rho_{HF1\text{-}4}{}^{c}$	$Q_{F}{}^{a}$ $C_{F}{}^{c}$ $T_{F}{}^{a}$ $T_{CW}{}^{a}$	$UA_{1\text{-}5}{}^{c}$ $UA_{C}{}^{c}$

[a]Operational data
[b]Computable from the available data
[c]Estimated from operational data

Only 12 groups have their values provided directly or computed from the existing on-line historical plant data. The rest are either unmeasured or impractical to measure; in particular the heat transfer coefficient UA and the solid composition C_F. To provide a reasonable estimation for these two variables, the steady state process model is validated against the plant data using a steady-state optimisation technique. Accordingly, the model is validated and the uncertain parameters are estimated by minimising the differences between the average plant data with the solution of the steady state model. The objective function for the validation Φ_V is as follows:

$$\min \Phi_V \tag{7.37}$$

$$\Phi_V = a_1 \times \left(\left(T_{P_1} - \tilde{T}_{P_1}\right)^2 + \left(T_{P_2} - \tilde{T}_{P_2}\right)^2 + \left(T_{P_3} - \tilde{T}_{P_3}\right)^2 + \left(T_{P_5} - \tilde{T}_{P_5}\right)^2 \right) + \cdots$$

$$a_2 \times \left(\rho_{P_5} - \tilde{\rho}_{P_5}\right)^2$$

s.t. $(\mathbf{7.1}) - (\mathbf{7.36})$ (steady state)

The wave symbols in (7.37) represent the average operational data. Different weighting coefficients $a_{1\text{-}2}$ are used to compare the relative importance of temperatures and product density.

The plant data for validation are $T_{P1\text{-}3}$, T_{P5} and ρ_{P5}, corresponding to four-effect operation, using the stages 1-3 and 5. The data have been sampled every 5 minutes for four hours operation. The profiles relative to the average values (0%) are shown in Figure 7.4.

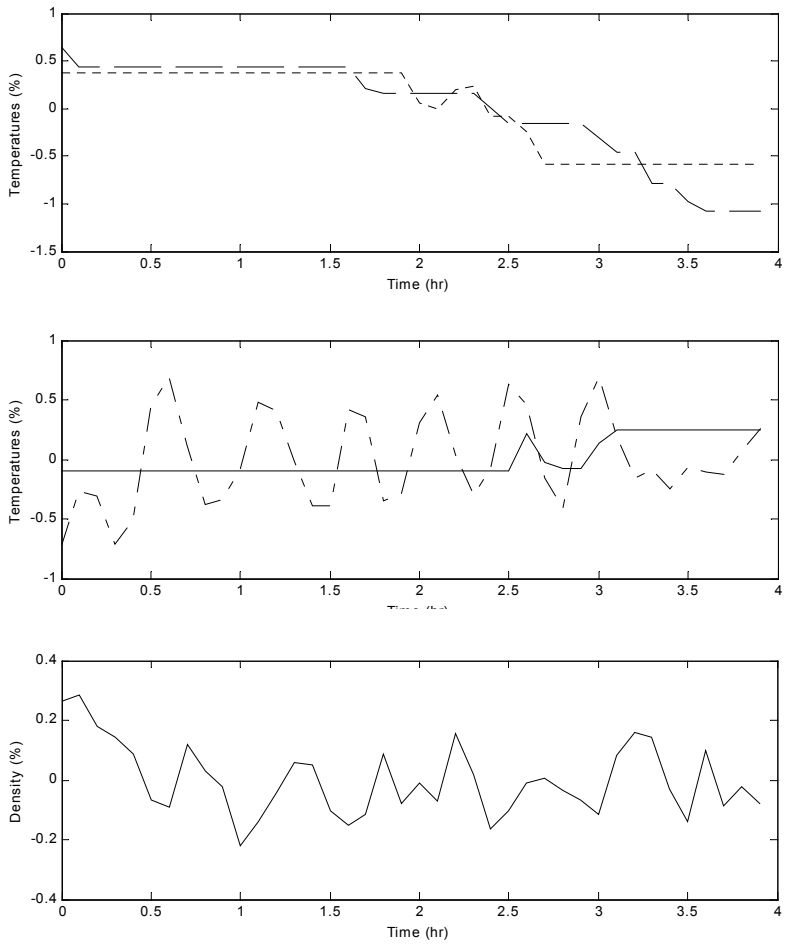


Figure 7.3 Plant Data Profiles for Model Validation, Relative to Average Values

\- - - - = Stage 1, —— = Stage 2, — - = Stage 3, —— = Stage 5

The validations have been slow to converge and sensitive to the initial conditions, due to strong state interactions. Despite this condition, the validation results are within 5% of the average plant data, as shown in Table 7.4. Based on this condition, the validation is considered acceptable. The estimated values of Q_{Ii}, C_F, ρ_{HFi}, M_V, UA_i and UA_c have been taken from the third validation, which give the smallest deviation from the plant average. The process parameters of stage 4 are assumed the same as the estimated values for stage 3.

Table 7.4 Validation Results

Parameters	Validation				Average
	1	2	3	4	
a_1	0.05	0.02	0.01	0.01	
a_2	10000	10000	10000	10000	
ρ_{P5}	-1.85%	-1.24%	-2.15%	-0.36%	-1.40%
T_{P1}	-2.08%	-1.87%	-1.94%	-1.49%	-1.85%
T_{P2}	2.73%	2.80%	2.72%	2.78%	2.76%
T_{P3}	0.69%	0.73%	0.67%	0.72%	0.70%
T_{P5}	-0.74%	-5.12%	-0.17%	-10.36%	-4.1%

7.6 Controllability Specifications

7.6.1 Controllability Objectives

As mentioned earlier, the expected liquor-burning evaporator performance measures is the maximum yield and the minimum energy costs, whilst maintaining the product temperature and compositions above target values, and other variables within upper and lower operational limits. In this case study, the expectation has been translated into the process economy Φ, which is the total amount of vaporisation minus the cost of heating and cooling energy (7.33).

The function is repeated for clarity:

$$\Phi = \sum_{i=1}^{5} M_{V_i} - \begin{pmatrix} 4\times10^{-4}\begin{pmatrix} 0.5 * M_{S_5}\left(H_{S_5} - E_{C_5}\right)+\cdots \\ M_{S_4}\left(H_{S_4} - E_{C_4}\right)+\ldots \\ M_{S_3}\left(H_{S_3} - E_{C_3}\right) \end{pmatrix}+\ldots \\ 1\times10^{-4}\times M_{CW}\left(E_{HW} - E_{CW}\right) \end{pmatrix}$$

The total vaporisation represent the total increases in solid composition, therefore the objective represents the process profit over the energy cost. This function is optimised subject to the given process dynamics, operational specifications, as well as disturbances and uncertainties.

The operational constraints are defined in their targeted values as follows:

$$\rho_{P_4} \geq \rho^{*}_{P_4} \tag{7.38}$$

$$\rho_{P_5} \geq \rho^{*}_{P_5} \tag{7.39}$$

$$T_{P_{5ss}} \geq T^{*}_{P_5} \tag{7.40}$$

$$T^{min}_{P_5} \leq T_{P_5} \leq T^{max}_{P_5} \tag{7.41}$$

$$h^{min}_{P_{1\text{-}5}} \leq h_{P_{1\text{-}5}} \leq h^{max}_{P_{1\text{-}5}} \tag{7.42}$$

$$0 \leq Q_{P_{1\text{-}5}} \leq 2Q^{max}_{P_{1\text{-}5}} \tag{7.43}$$

$$Q^{min}_{I_{1\text{-}4}} \leq Q_{I_{1\text{-}4}} \leq Q^{max}_{I_{1\text{-}4}} \tag{7.44}$$

$$0 \leq M_{V_{1\text{-}5}} \leq M^{max}_{V_{1\text{-}5}} \tag{7.45}$$

$$0 \leq M_{S_{1\text{-}5}} \leq M^{max}_{S_{1\text{-}5}} \tag{7.46}$$

$$M^{min}_{CW} \leq M_{CW} \leq M^{max}_{CW} \tag{7.47}$$

The above specifications are applied for the dynamic profile, except equation (7.40), which is applied for the calculation of initial condition of T_{P5} (T_{P5ss}).

The objective is evaluated across the possible process and controller structure within the superstructure, with the specification given in the following sections.

7.6.2 Process Structure Specifications

The process structure problem is to decide the number of effects to operate in satisfying the profit objective. This decision also affects the consequences of steam and vapor distribution within the process. The live steam can only be supplied to stages 3-5, therefore the possible number of effects is 3-5. As mentioned earlier, at least one of the stages 3 or 4 will operate. If both stages are operating, vapor from both stages will mix. Furthermore, only steam flowing to stage 4 can be manipulated, while the flow to stage 3 is fixed during the process operation. To capture these possibilities within the superstructure, the binary design variables $y_{P3\text{-}5}$ are assigned to the operating amounts of steam, vaporisation, and recycles. These assignments ensure that the unoperational stage is not receiving any steam, producing any vapor, recycling the liquor, and not exchanging heat.

These are translated to the process structure specifications in (7.48)-(7.49), the modification of the operational specification (7.43)-(7.46) into (7.50)-(7.52), and the process model (7.9), (7.10), (7.26), and (7.27) into (7.53)-(7.56) as follows:

$$y_{P_{3\text{-}5}} \in \{0,1\} \tag{7.48}$$

$$y_{P_3} + y_{P4} \geq 1 \tag{7.49}$$

$$Q_{I_i}^{min} \times y_{P_i} \leq Q_{I_i} \leq Q_{I_i}^{max} \times y_{P_i} \quad i = 3,4 \tag{7.50}$$

$$0 \leq M_{V_i} \leq M_{V_i}^{max} \times y_{P_i} \qquad i = 3,4,5 \tag{7.51}$$

$$0 \leq M_{S_i} \leq M_{S_i}^{max} \times y_{P_i} \qquad i = 3,4,5 \tag{7.52}$$

$$M_{V_3} = M_{V34}\left(y_{P_4} H_{V_4} - H_{V34}\right) / \left(y_{P_4} H_{V_4} - y_{P_3} H_{V_3}\right) \tag{7.53}$$

$$M_{V_4} = M_{V34}\left(y_{P_3} H_{V_3} - H_{V34}\right) / \left(y_{P_3} H_{V_3} - y_{P_4} H_{V_4}\right) \tag{7.54}$$

$$y_{Pi} M_{S_i}\left(H_{S_i} - E_{C_i}\right) = M_{HF_i}\left(E_{HD_i} - E_{HF_i}\right) \qquad i = 3,4,5 \tag{7.55}$$

$$T_{HD_i} = T_{S_i} - \left(T_{S_i} - T_{HF_i}\right)\left(1 - \exp\left(\frac{-UA_i y_{Pi}}{M_{HF_i} cp_{HF_i}}\right)\right) \qquad i = 3,4,5 \tag{7.56}$$

Applying the above specification into the steady state model would force an un-operational stage to adopt the steady state condition of the preceding stage. For instance, if stage 3 is not operating, the steady state values of level, composition ratio, temperature, density, enthalpy and product flow-rate of this stage become identical to those of stage 2. Then, the computation for stage 4 can be performed using the values from stage 3 as usual.

To ensure that the above conditions are extended to the corresponding dynamic profiles, the un-operational derivative terms at the corresponding DAE are further modified to adopt the derivative terms of the preceding stage. For this purpose, the flash tank equations (7.28)-(7.30) are modified for stages 3-5 as in (7.57)-(7.62).

Flash tanks: i = 3, 4

$$\frac{d(h_{P_i})}{dt} = \left(1 - y_{P_i}\right)\frac{d(h_{P_{i-1}})}{dt} + y_{P_i}\begin{pmatrix} Q_{HF_i}\rho_{HF_i} - \cdots \\ \left(Q_{I_i} + Q_{P_i}\right)\rho_{P_i} - M_{V_i} \end{pmatrix} / A_i\rho_{P_i} \tag{7.57}$$

$$\frac{d(x_{Pi})}{dt} = \left(1 - y_{P_i}\right)\frac{dx_{Pi-1}}{dt} + y_{P_i}\begin{pmatrix} Q_{HF_i}x_{HF_i} - \cdots \\ \left(Q_{I_i} + Q_{P_i}\right)x_{P_i} \end{pmatrix} / A_i h_{P_i} \tag{7.58}$$

$$\frac{d\left(E_{P_i}\right)}{dt} = \left(1 - y_{P_i}\right)\frac{dE_{P_{i-1}}}{dt} + y_{P_i}\begin{pmatrix} Q_{P_{i-1}}\rho_{P_{i-1}}\left(E_{P_{i-1}} - E_{P_i}\right)\cdots \\ -M_{V_i}\left(H_{V_i} - E_{P_i}\right)\cdots \\ +M_{S_i}\left(H_{S_i} - E_{C_i}\right) \end{pmatrix} / A_i\rho_{P_i}h_{P_i} \tag{7.59}$$

Flash tank 5

$$\frac{d(h_{P_5})}{dt} = \left(1 - y_{P_5}\right)\frac{d(h_{P_4})}{dt} + y_{P_5}\begin{pmatrix} Q_{HF_5}\rho_{HF_5} - \cdots \\ \left(Q_{I_5} + Q_{P_5}\right)\rho_{P_5} - M_{V_5} \end{pmatrix} / A_5\rho_{P_5} \tag{7.60}$$

$$\frac{d\left(x_{P_5}\right)}{dt} = \left(1 - y_{P_5}\right)\frac{dx_{P_4}}{dt} + y_{P_5}\begin{pmatrix} Q_{HF_5}x_{HF_5} - \cdots \\ \left(Q_{I_5} + Q_{P_5}\right)x_{P_5} \end{pmatrix} / A_5 h_{P_5} \tag{7.61}$$

$$\frac{d\left(E_{P_5}\right)}{dt} = \left(1 - y_{P_5}\right)\frac{dE_{P_4}}{dt} + y_{P_5}\begin{pmatrix} Q_{P_4}\rho_{P_4}\left(E_{P_4} - E_{P_5}\right)\cdots \\ -M_{V_5}\left(H_{V_5} - E_{P_5}\right)\cdots \\ +M_{S_5}\left(H_{S_5} - E_{C_5}\right) \end{pmatrix} / A_5\rho_{P_5}h_{P_5} \tag{7.62}$$

These specifications accommodate the selection of a four or five-effects operation, with six alternative configurations ([y_{P3} y_{P4} y_{P5}] = [1 0 0], [0 1 0], [1 1 0], [1 0 1], [0 1 1], [1 1 1]) at both the steady state and the DAE model.

7.6.3 Controller Structure Specification

The controller structure problem is to select the best matching between the manipulated and the output variables in feedback and feedforward control loops, to maintain the dynamics within the specified limits and provide maximum economy. Table 7.2 shows the vast variety of possible control loops, which would be assessed systematically within the superstructure in this study.

The dynamics under study are the level inventories h_{Pi}, the product densities at stages 1, 4, and 5 $\rho_{1,4\text{-}5}$, and the product temperature T_{P5}. The levels are open-loop unstable, while the density and temperature of the product are self-regulated. However, all are strongly interactive (Kam and Tade, 2000). Based on Table 7.2, the candidate of manipulated variables for feedback loops are the product flow-rates Q_{Pi}, the cooling water flow-rate M_{CW}, the steam supplies M_{V34} and M_{V5}, and the final vaporisation M_{V5}. Flow-rates Q_{Pi} are the favorite candidates to manipulate levels, for their direct effects (7.28), (7.57), (7.59). However, they also directly affect the densities (7.29), (7.58), (7.60). It would be interesting to evaluate whether the interaction would degrade the density when integral actions are applied on h_{Pi}-Q_{Pi}.

Several possible feedback and feedforward control loops relating the output and manipulated variables have been supplied to process superstructure. To keep the problem in a manageable size, the feedback loops have been pre-designed for stability in Proportional – Integral (PI) control scheme, defined by the gains $K_{C1\text{-}12}$ and reset times $\tau_{1\text{-}12}$. The feedforward loops are in Proportional only scheme. There are 12 control loops considered. The binary variables $y_{C1\text{-}12}$ are attached to each loop and relevant constraint. The active loop can be tuned with scaling variables $0 \leq \alpha_{1\text{-}12} \leq 2$.

Since the levels are open-loop unstable, the proportional gains are active in all cases. The integral actions, however, can be deactivated using y_{c1}=0. For stages 1-4, the product flow-rates $Q_{P1\text{-}4}$ are the manipulated variables. At the final stage, the level may be manipulated using either Q_{P5}, or M_{V5}. The product density ρ_{P5} should be maintained above the target level. To meet this goal, ρ_{P5} can be controlled using M_{S5}, M_{V5} or Q_{P5}. The control loops ρ_{P1}-M_{CW} and ρ_{P4}-M_{S34} are also evaluated in their role of achieving the goal $\rho_{P5} \geq \rho_{P5}^*$. The operating value of the product temperature T_{P5ss} should be maintained above the target value, but the dynamics are allowed within the lower and upper limits. For this purpose, T_{P5} may be controlled using M_{S5}, or M_{V5}. These controller specifications are listed in Table 7.5.

Table 7.5 Control Loops and Parameters

Integer variable	Output variable	Manipulated variable	Kci	τ_i (hr)	Feedback / feedforward	Note
-	$h_{P1\text{-}4}$	$Q_{P1\text{-}4}$	25		Feedback	Always active
y_{C1}	$h_{P1\text{-}4}$	$Q_{P1\text{-}4}$		1	Feedback	Activating
y_{C1}	h_{P5}	Q_{P5}		1	Feedback	integral action
y_{C1}	h_{P5}	M_{V5}		0.5	Feedback	only
y_{C2}	h_{P5}	Q_{P5}	25		Feedback	Activating
y_{C3}	h_{P5}	M_{V5}	20		Feedback	proportional gain
y_{C4}	ρ_{P1}	M_{CW}	-5000	0.1	Feedback	
y_{C5}	ρ_{P5}	Q_{P5}	50	0.05	Feedback	
y_{C6}	ρ_{P4}	M_{S34}	-30	2	Feedback	Activating
y_{C7}	ρ_{P5}	M_{S5}	100	1	Feedback	both
y_{C8}	ρ_{P5}	M_{V5}	-100	1	Feedback	proportional gain
y_{C9}	T_{P5}	M_{S5}	-0.2	1	Feedback	and
y_{C10}	T_{P5}	M_{V5}	-0.01	0.1	Feedback	integral action
y_{C11}	Q_F	M_{S34}	0.1		Feedforward	
y_{C12}	Q_F	M_{CW}	100		Feedforward	

Only a multi-loop control scheme is considered in this study, therefore every manipulated variable can only be used once. The controllers are also activated based on the process structure. For instance, if the final stage is un-operational, then no controllers in this stage can be activated. The constraints defining these controller structure specifications are as follows:

$$y_{C_{1\text{-}12}} \in \{0,1\} \qquad i = 1,2,\ldots,12 \tag{7.63}$$

$$y_{C_2} + y_{C_3} = y_{P_5} \tag{7.64}$$

$$y_{C_4} + y_{C12} = 1 \tag{7.65}$$

$$y_{C_{10}} + y_{C11} = 1 \tag{7.66}$$

$$y_{C_7} + y_{C_8} = y_{C2} \tag{7.67}$$

$$y_{C_5} = y_{C3} \tag{7.68}$$

$$y_{C_{10}} = y_{C7} \tag{7.69}$$

$$y_{C_5} + y_{C_8} = y_{C9} \tag{7.70}$$

$$y_{C_2} + y_{C5} \le y_{P5} \tag{7.71}$$

$$y_{C_7} + y_{C9} \le y_{P5} \tag{7.72}$$

$$y_{C_3} + y_{C8} + y_{C10} \le y_{P5} \tag{7.73}$$

$$y_{C_8} + y_{C5} + y_{C7} \le y_{P5} \tag{7.74}$$

$$y_{C_9} + y_{C10} \le y_{P5} \tag{7.75}$$

$$y_{C_6} \le y_{P_4} \tag{7.76}$$

$$0 \le \alpha_i \le 2yc_i \qquad i = 1,2,\ldots,12 \tag{7.77}$$

While in total there are 15 binary variables, the above specifications have reduced the possible combinations to 36.

The corresponding control equations within the DAE are as follows:

$$e_{\rho_{P_i}} = \rho_{Pi} - \rho_{Piss}, \quad \frac{dI_{\rho_{Pi}}}{dt} = e_{\rho_{Pi}}, \qquad i = 1,4,5 \tag{7.78}$$

$$e_{h_{Pi}} = h_{P_i} - h_{P_{iss}}, \quad \frac{dI_{h_{P_i}}}{dt} = e_{h_{P_i}}, \quad i = 1,2,3,4,5 \tag{7.79}$$

$$e_{TP5} = T_{P_5} - T_{P_{5ss}}, \quad \frac{dI_{T_{P_5}}}{dt} = e_{T_{P_5}} \tag{7.80}$$

$$Q_{P_i} = Q_{P_{iss}} - K_{C1}\left(e_{h_{P_i}} + \alpha_1 \frac{I_{h_{P_i}}}{\tau_1} \right), \qquad i = 1,2,3,4 \tag{7.81}$$

$$Q_{P_5} = Q_{P_{5ss}} - \alpha_2 K_{C1}\left(e_{h_{P_i}} + \alpha_1 \frac{I_{h_{P_i}}}{\tau_1} \right) + \alpha_5 K_{C5}\left(e_{\rho_{P_5}} + \frac{I_{\rho_{P_5}}}{\tau_5} \right) \tag{7.82}$$

$$M_{V_5} = M_{V_{5ss}} - \alpha_3 K_{C3}\left(e_{h_{P_5}} + \alpha_1 \frac{I_{h_{P_5}}}{\tau_3} \right) + \alpha_8 K_{C8}\left(e_{\rho_{P_5}} + \frac{I_{\rho_{P_5}}}{\tau_8} \right) + \ldots + \alpha_{10} K_{C10}\left(e_{T_{P_5}} + \frac{I_{T_{P_5}}}{\tau_{10}} \right) \tag{7.83}$$

$$M_{S_5} = M_{S_{5ss}} - \alpha_7 K_{C7}\left(e_{\rho_{P_5}} + \frac{I_{\rho_{P_5}}}{\tau_7} \right) + \alpha_9 K_{C9}\left(e_{T_{P_5}} + \frac{I_{T_{P_5}}}{\tau_9} \right) \tag{7.84}$$

$$M_{S_{34}} = M_{S_{34ss}} - \alpha_6 K_{C6}\left(e_{\rho_{P_4}} + \frac{I_{\rho_{P_4}}}{\tau_6} \right) + \alpha_{11} K_{C11}\left(Q_F - Q_{F_{ss}} \right) \tag{7.85}$$

$$M_{CW} = M_{CW} - \alpha_4 K_{C4}\left(e_{\rho_{P_1}} + \frac{I_{\rho_{P_1}}}{\tau_4} \right) + \alpha_{12} K_{C12}\left(Q_F - Q_{F_{ss}} \right) \tag{7.86}$$

$$M_{S_4} = M_{S_{34}} - y_{P_3} M_{S_{3ss}} \tag{7.87}$$

$$M_{S_3} = \left(1 - y_{P_4}\right) M_{S_{34}} - y_{P_4} M_{S_{3ss}} \tag{7.88}$$

7.7 Controllability Assessments

The controllability assessments are performed subject to two cases of disturbance consideration. The first case is the worst-case analysis, where the disturbances are assumed as sets of step functions, with magnitudes varying within the upper and lower bounds. The assessment is performed using the non-sequential framework. The second case considers the general disturbance profiles based on the plant data, and assessed using the sequential version. The assessments are programmed in MATLAB and operated in a Pentium IV PC 2.4 GHz with 256 MB RAM. The results are discussed in the following sections.

7.7.1 The Worst-case Disturbance Characterisation

In this case, the disturbances are the feed flow-rate Q_F, feed temperature T_F, feed composition x_F, and cooling water temperature T_{CW}. They are assumed as sets of the step functions. The magnitudes are assumed varying within the upper and lower bounds defined by the percent of deviation from the nominal conditions, and their values are given in Table 7.6. The lower and upper bounds of Q_F, T_F, and T_{CW} are the taken and rounded from the worst deviation of the available historical plant data. While the feed composition x_F is a very likely disturbance, its variations are difficult to measure. Therefore, there is no historical plant data for this variable. In this study, this variable is assumed to vary around the nominal value found from the validation step described at section 7.5.3.

Table 7.6 Worst-Case Disturbance Ranges

Parameters	Lower bound (deviation)	Upper bound (deviation)
Q_F	-25%	25%
T_F	-1.5%	0.5%
x_F	-5%	5%
T_{CW}	-10%	5%

This assessment involves 75 variables, 136 steady state equations, 154 DAE, 103 inequality steady state constraints, 31 dynamic path constraints and 15 logic propositions. Due to the size of the problem, the assessment takes 1198 CPU minutes (19.96 hours) to complete. The results are given in Table 7.7.

.

Table 7.7 Controllability Assessment Results

	Steady State Optimum	Closed-loop optimum	w	u
Φ	9.12	9.05		
ρ_{P4}	$1.03\rho_{P4}$*	$1.03\rho_{P4}$*		
ρ_{P5}	$1.03\rho_{P5}$*	$1.02\rho_{P5}$*		
T_{P5}	T_{P5}*	$1.03T_{P5}$*		
Q_{I1}	$1.41Q_{I1}^{min}$	$1.41Q_{I1}^{min}$		
Q_{I2}	$1.38Q_{I2}^{min}$	$1.38Q_{I2}^{min}$		
Q_{I3}	0.000	0.000		
Q_{I4}	$1.30Q_{I4}^{min}$	$1.30Q_{I4}^{min}$		
y_{P3} (m_{S3})	0 (0.000)	0 (0.000)		
y_{P4} (m_{S4})	1 (93.4% steam)	1 (91.8% steam)		
y_{P5} (m_{S5})	1 (6.5% steam)	1 (8.2% steam)		
y_{c1} (α_1)	1 (1.000)	1 (1.000)		
y_{c2} (α_2)	1 (1.000)	1 (1.000)	h_{P5}	Q_{P5}
y_{c3} (α_3)	0 (0.000)	0 (0.000)	h_{P5}	M_{V5}
y_{c4} (α_4)	0 (0.000)	0 (0.000)	ρ_{P1}	M_{CW}
y_{c5} (α_5)	0 (0.000)	0 (0.000)	ρ_{P5}	Q_{P5}
y_{c6} (α_6)	1 (1.000)	0 (0.000)	ρ_{P4}	M_{S34}
y_{c7} (α_7)	0 (0.000)	0 (0.000)	ρ_{P5}	M_{S5}
y_{c8} (α_8)	0 (0.000)	0 (0.000)	ρ_{P5}	M_{V5}
y_{c9} (α_9)	0 (0.000)	0 (0.000)	T_{P5}	M_{S5}
y_{c10} (α_{10})	1 (1.000)	1 (1.000)	T_{P5}	M_{V5}
y_{c11} (α_{11})	0 (0.000)	0 (0.000)	Q_F	M_{S34}
y_{c12} (α_{12})	1 (1.000)	0 (0.000)	Q_F	M_{CW}
Variable groups	$[h_{P1-5}, \rho_{P4}, T_{P5}]$ $[Q_{P1-4,} M_{S34,} M_{CW}, M_{V5}]$ $[Q_{P5}, \rho_{P5}]$	$[h_{P1-5}, T_{P5}]$ $[Q_{P1-4,} \rho_{P4}, Q_{P5}, \rho_{P5}, M_{V5}]$		
$\overline{w}$	h_{P1}, M_{S34}, Q_{P5}	T_{P5}, Q_{P5}		
θ^k	$[Q_F^{min}, T_F^{max}, x_F^{max}, T_{CW}^{min}]$, $[Q_F^{max}, T_F^{min}, x_F^{min}, T_{CW}^{max}]$	$[Q_F^{min}, T_F^{max}, x_F^{max}, T_{CW}^{min}]$, $[Q_F^{max}, T_F^{min}, x_F^{min}, T_{CW}^{max}]$		
r-OCI	0.708	1		
PVAR	186.708	81.393		

This assessment converges in two major iterations, each involving the outer and inner levels. The first outer level is the steady state MINLP problem based on nominal disturbance values. The problem raises an optimisation tree containing an NLP sub problem at each branch, which is evaluated using the modified Branch and Bound algorithm developed in chapter 5. It takes 22 branch evaluations and 2.32 minutes to complete this level. The optimal steady state process structure found is the four-effect operations using the stages 1, 2, 4 and 5 ($[y_{P3}\ y_{P4}\ y_{P5}] = [0\ 1\ 1]$). The logic proposition level suggests PI control for all levels control loops $h_{P1\text{-}4}$-$Q_{P1\text{-}4}$, as well as the loops ρ_{P4}-M_{S34}, T_{P5}-M_{V5}, and Q_F-M_{CW} as tentative controller structures.

The strong interaction between variables and the existing of binary variables within the process model, leads to the local solution of the problem, and the high computational cost, as indicated by the steady state solutions in the first outer level. The problem gets even bigger when the process dynamics need to be solved, starting from the first inner level. In this level, the feasibility of the closed loop performances over a 24 hour time horizon subject to all disturbance combinations is evaluated. Each DAE requires 3.2 seconds on average to complete. To keep the problem to a manageable size, the disturbance combinations are considered at their extreme values. Therefore, the feasibility test at this level is subject to 16 disturbance combinations.

The feasible output space is defined by the measured output variables $h_{P1\text{-}5}$, $\rho_{P1,4,5}$, T_{P5}, and the manipulated variables $Q_{P1\text{-}5}$, M_{V5}, M_{S34}, M_{S5}, and M_{CW}. The redundancy analysis finds three collinear groups and selects the corresponding functional variables $\overline{w}$ as shown in Table 7.7. The analysis subsequently determines two critical disturbance combinations θ^k. The steam M_{S5} is not used at this level, and accordingly it does not appear in $\overline{w}$.

The process dynamics are infeasible (r-OCI = 0.708) with a total variation in the process economy of 108.708, which is 49.2% of the nominal process economy over a 24 hours operational period. This infeasibility is due to the manipulation of M_{S34}, which is not only an additional cost to the process economy, but also unrealisable in

practice. Therefore, the second outer level would assess whether the loop ρ_{P34}-M_{S34} can be tuned to resume the process feasibility, or be replaced with another alternatives.

The second outer level solves the dynamic MINLP problem associated with the process and controller structure synthesis. This time, the subproblems are dynamic multi-period NLP. Each subproblem requires the solution of a steady state process model for nominal conditions and multiple sets of DAE, as many as the number of disturbance combinations found by the preceding redundancy analysis. Due to the accumulated interaction and high index DAE, one subproblem can take 90 – 120 minutes to converge, which gives the overall 19.96 hours completion time.

The result of the second outer level maintains the four-effect operation ($[y_{P3}\ y_{P4}\ y_{P5}] = [0\ 1\ 1]$) and the design values Q_{Ii}. The controller structures are nevertheless changed. While PI level control is still recommended ($y_{C1} = 0$), the only other loop is T_{P5}-M_{V5}. The corresponding redundancy analysis selects Q_{P5} and T_{P5} as functional variables. The steam flow M_{S34} and M_{S5} are not used in this structure, therefore they are not considered in the analysis. The dynamic performances due to the worst disturbance combinations for the solution of both the first and second levels are given in Figures 7.4 – 7.11.

These results highlight the energy cost minimisation of the given process economy Φ (7.33). Activating the loop ρ_{P4}-M_{S34} introduces a large fluctuation in M_{S34}, which propagates back to the preceding stages and forces the process dynamics further away from the feasible conditions. These are particularly shown in Figure 7.5. Therefore, M_{S34} is better not functioned as manipulated variables. However, they still has a strong role as design variables, which are intended to determine the higher initial values for the densities $\rho_{P4\text{-}5}$, such that their regulatory dynamics are easier maintained within the feasible operating region.

The same case also applies to M_{S5}. The super-concentrator at the final stage indeed has the co-current configuration that breaks the interaction between the product and the previous stages. Any adjustment at this stage would not be recycled to the previous

stages, therefore preventing unnecessary oscillations. However, manipulation using M_{S5} is not recommended for the given controllability objective, because the quantity is very small and easily exhausted when used for control purposes. Furthermore, as already shown in the M_{S34} case, the manipulations would be at infeasible operational region.

Therefore, the remaining available manipulated variables at this stage are Q_{P5} and M_{V5}. The usage of M_{V5} for h_{P5} control is possible, but it significantly drops the temperature T_{P5} below the allowable limits. The possible compensation is to use the loop T_{P5}-M_{S5}. However, this strategy requires a delicate balance between M_{V5} and M_{S5}. M_{S5} should have at least the same amount as M_{S34} to maintain feasible manipulation. Otherwise, it leads to a high-energy cost, which is not in favor with the process economy $\mathbf{\Phi}$.

Therefore, Q_{P5}-h_{P5} and T_{P5}-M_{V5} are the best control options for the super-concentrator, and this strategy is indeed found in the second outer level. This temperature control is relatively easy to implement, both computationally and in practice. The fluctuation in process economy is 37% of the total nominal, and the profile is shown in Figure 7.11. The integral actions in this level of controls do not significantly affect the density in comparison to the usage of live steam as manipulated variables. Figure 7.8 shows that M_{S4} and M_{S5} are not used as manipulated variable. Therefore, there is no dynamics on them. Overall, the control strategy found in the second outer level, which is all PI for $h_{P1\text{-}2,\ 4\text{-}5}$–$Q_{P1\text{-}2,\ 4\text{-}5,}$ and T_{P5}-M_{V5}, is the recommended control strategy for the process.

One last comment on the cooling water flowrate M_{CW} is that it is too slow and indirect for manipulating level h_{P1} or density ρ_{P1}. It does not provide significant assistance for process controllability. Therefore, it is not recommended as manipulated variables.

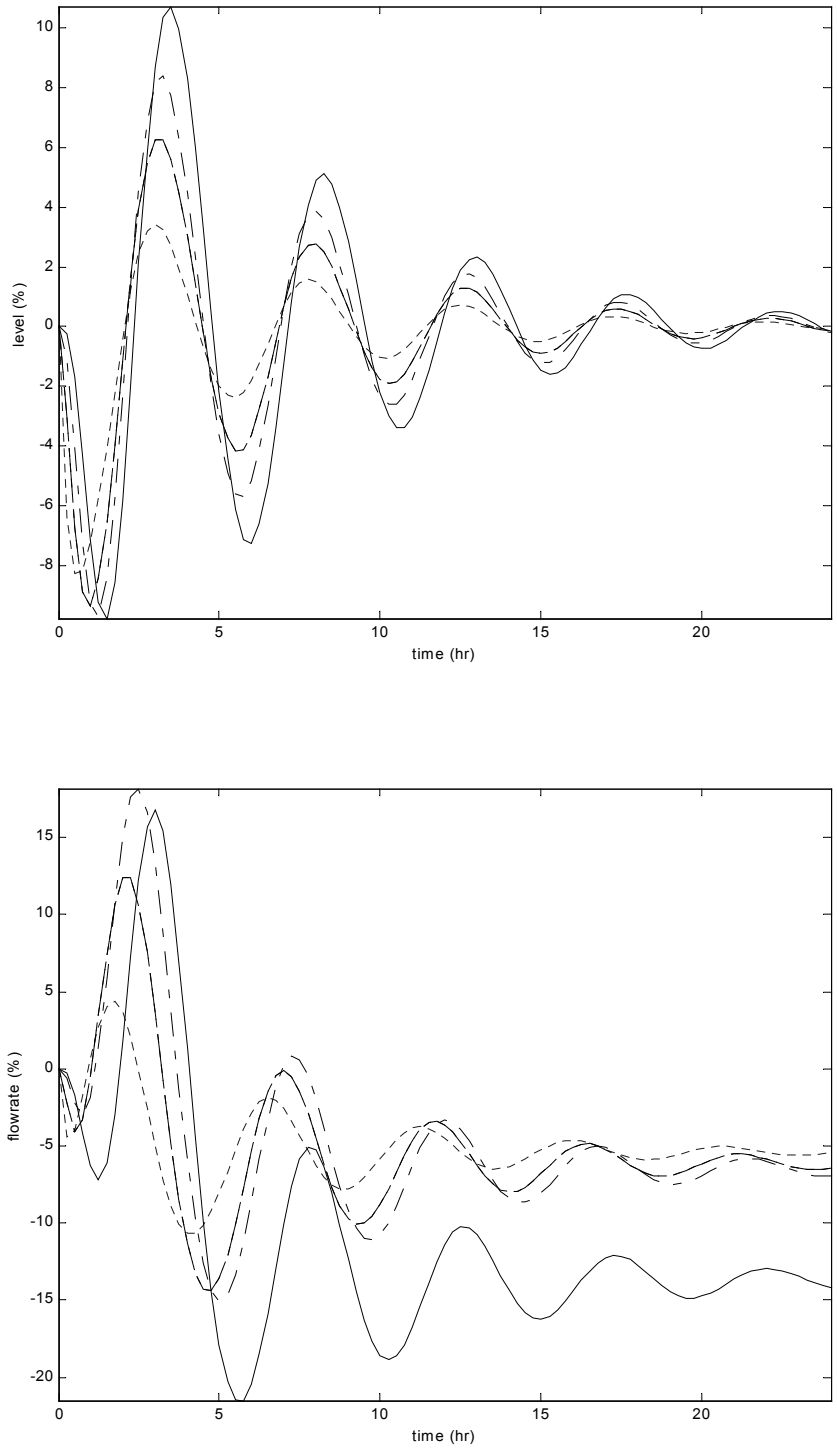


Figure 7.4 Worst-Case Analysis: Level and Flowrate Profiles in the First Inner-Level

---- = Stage 1, — — = Stage 2, — - — = Stage 4, — = Stage 5

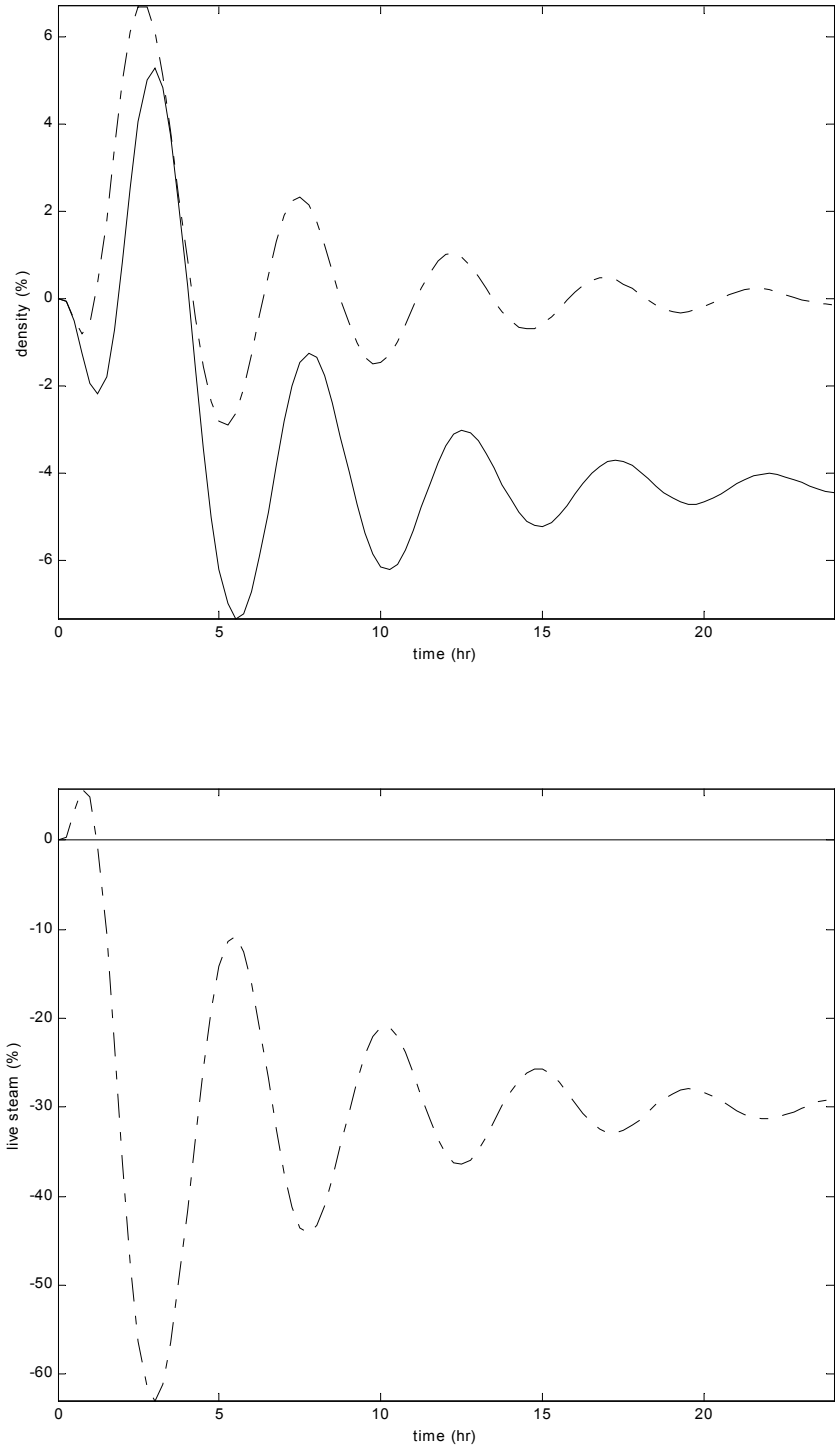


Figure 7.5 Worst-Case Analysis: Density and Steam Profiles in the First Inner-Level

— - — = Stage 4, — = Stage 5

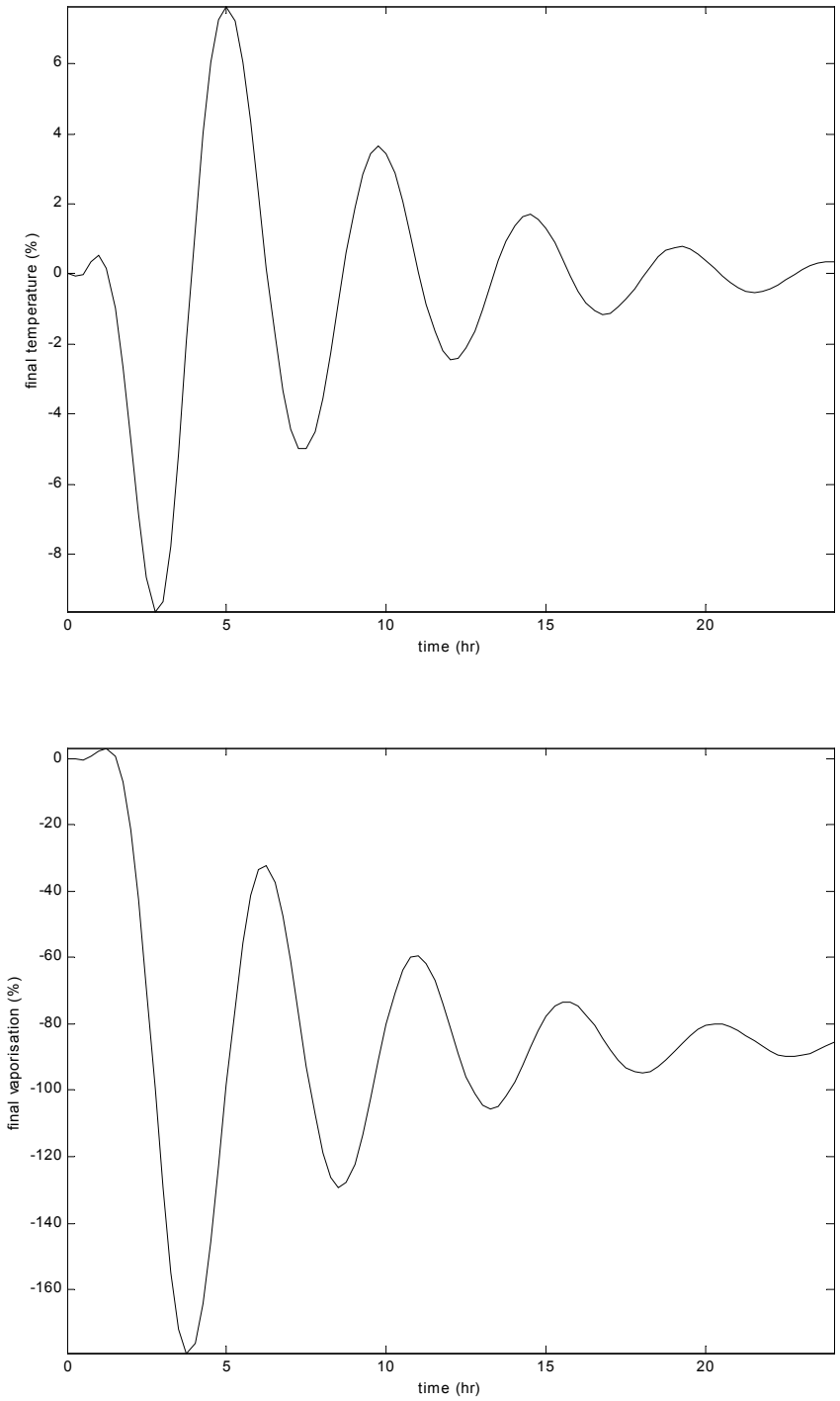


Figure 7.6 Worst-Case Analysis: Temperature and Vaporisation Profiles in the First Inner-Level

— = Stage 5

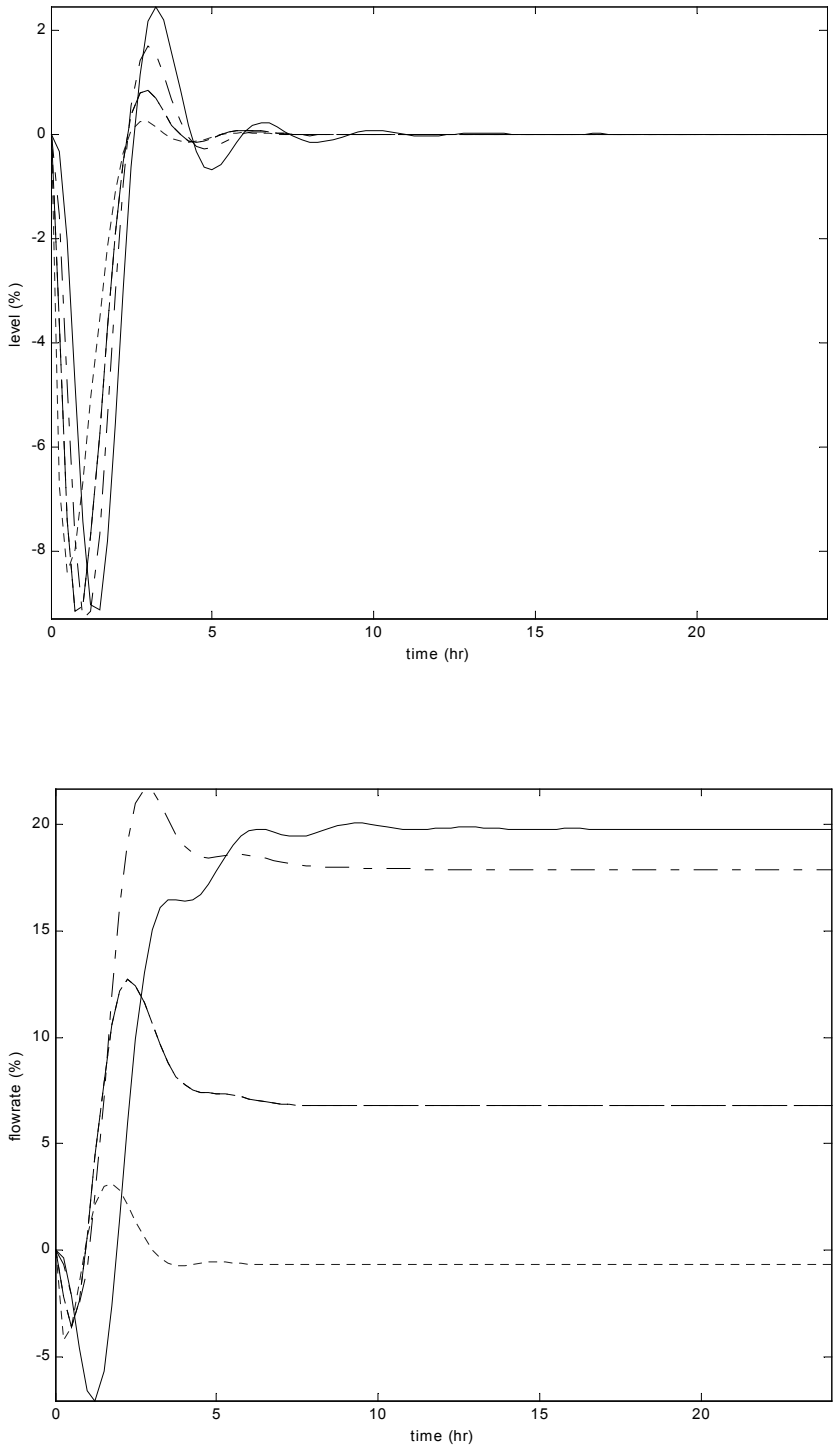


Figure 7.7 Worst-Case Analysis: Level and Product Flowrate Profiles in the Second Inner-Level

---- = Stage 1, — — = Stage 2, — - — = Stage 4, — = Stage 5

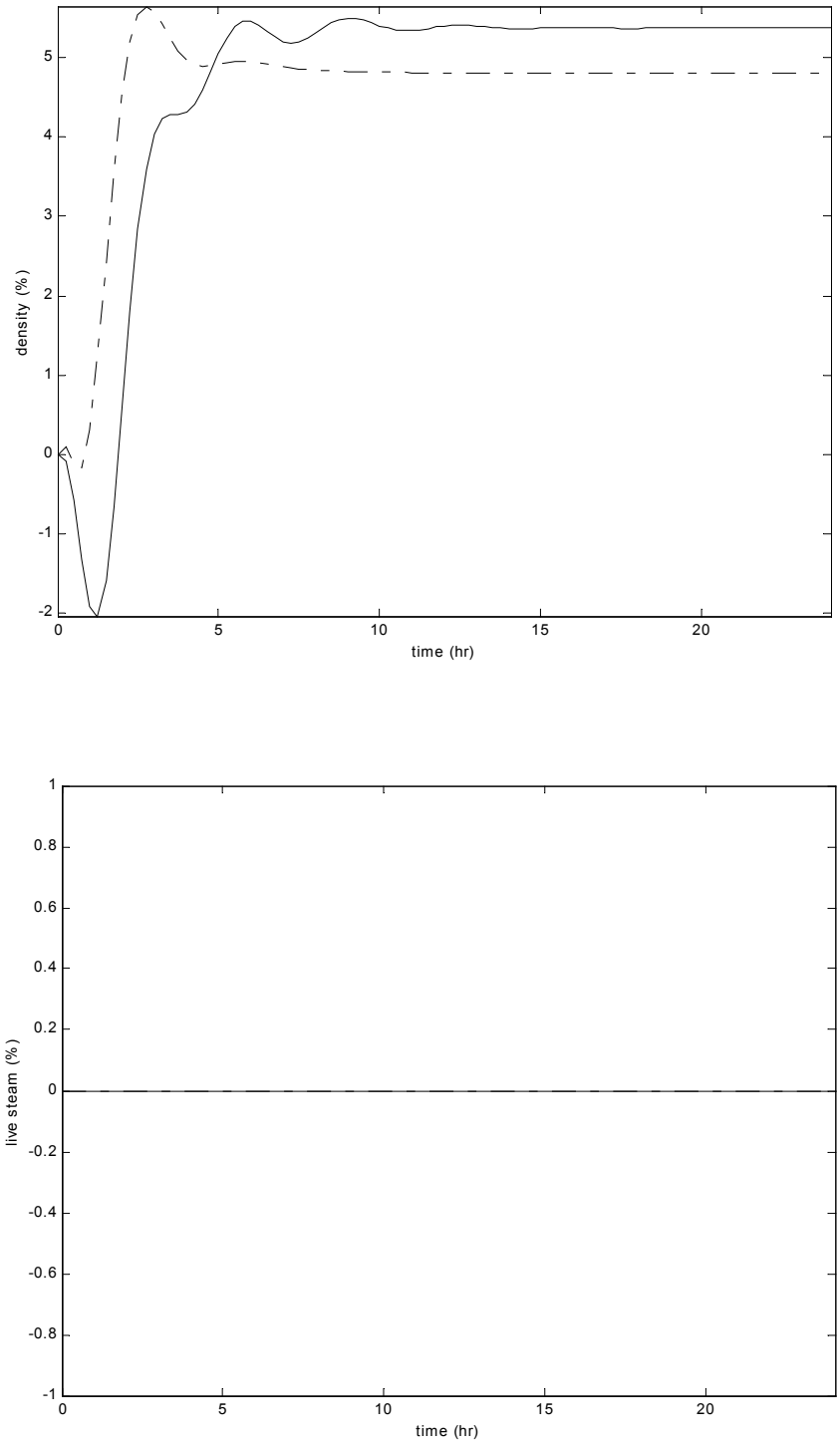


Figure 7.8 Worst-Case Analysis: Density and Live Steam Profiles in the Second Inner-Level

— - — = Stage 4, — = Stage 5

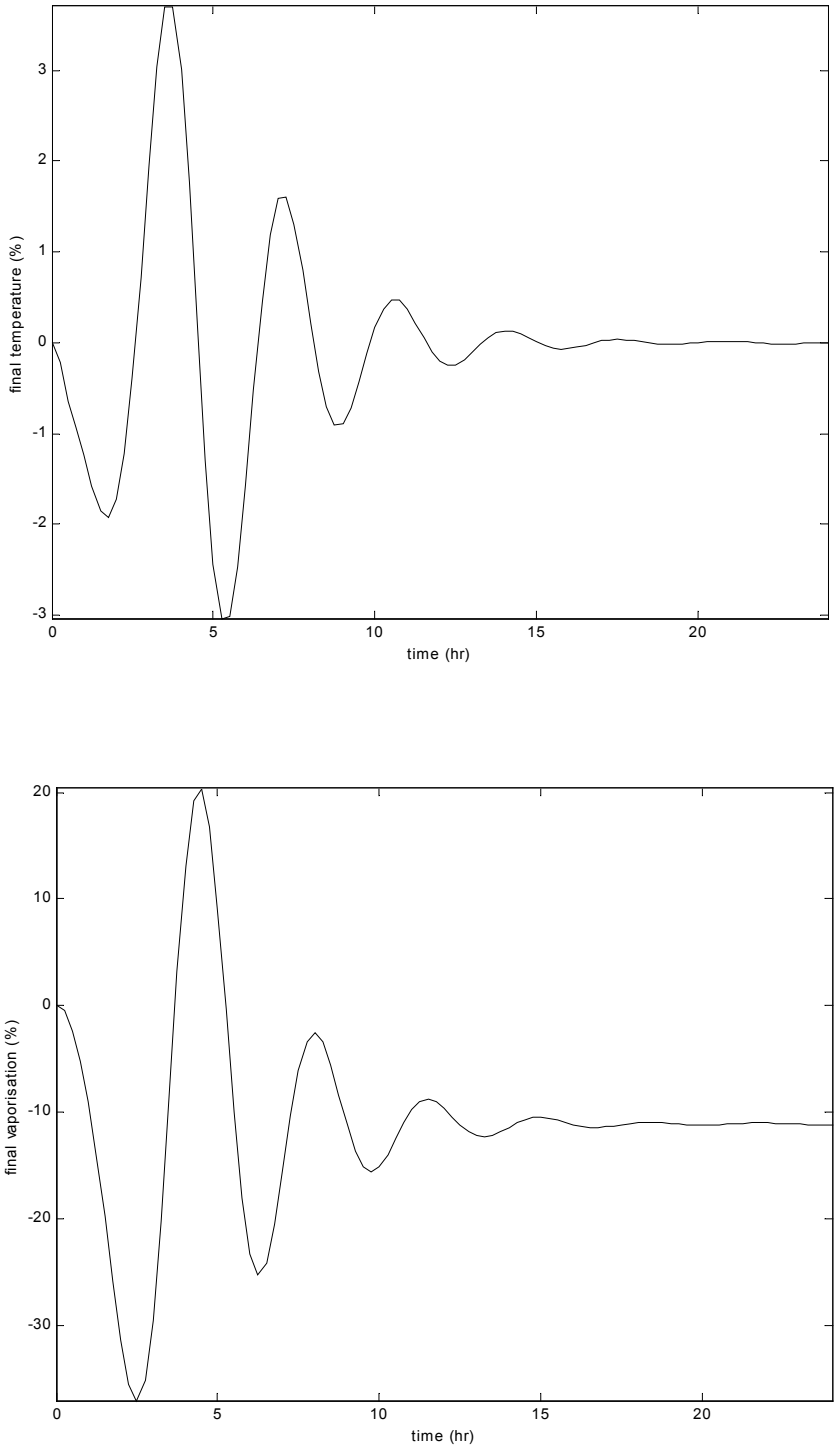


Figure 7.9 Worst-Case Analysis: Temperature and Vaporisation Profiles in the Second Inner Level

— = Stage 5

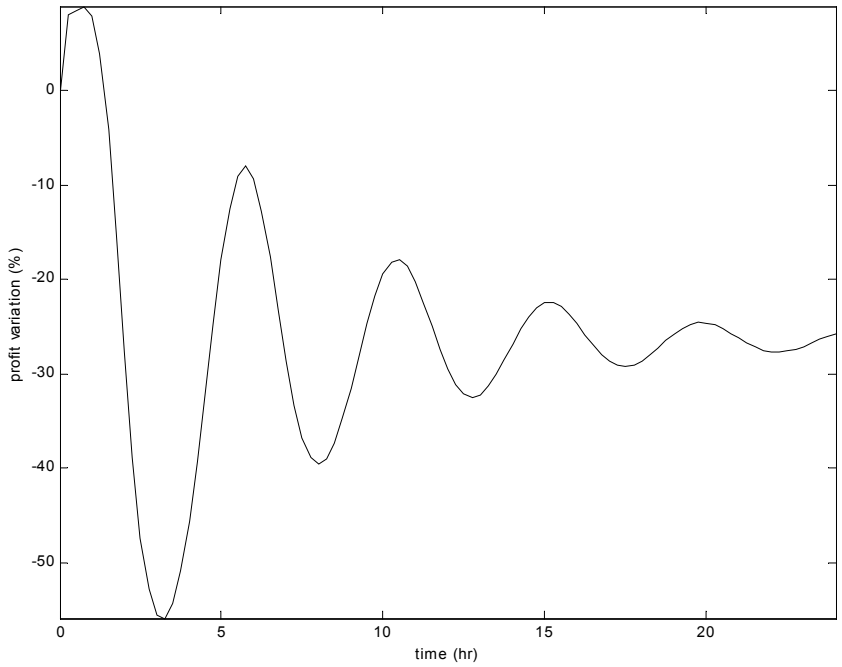


Figure 7.10 Worst case analysis: profit profiles at the first inner level

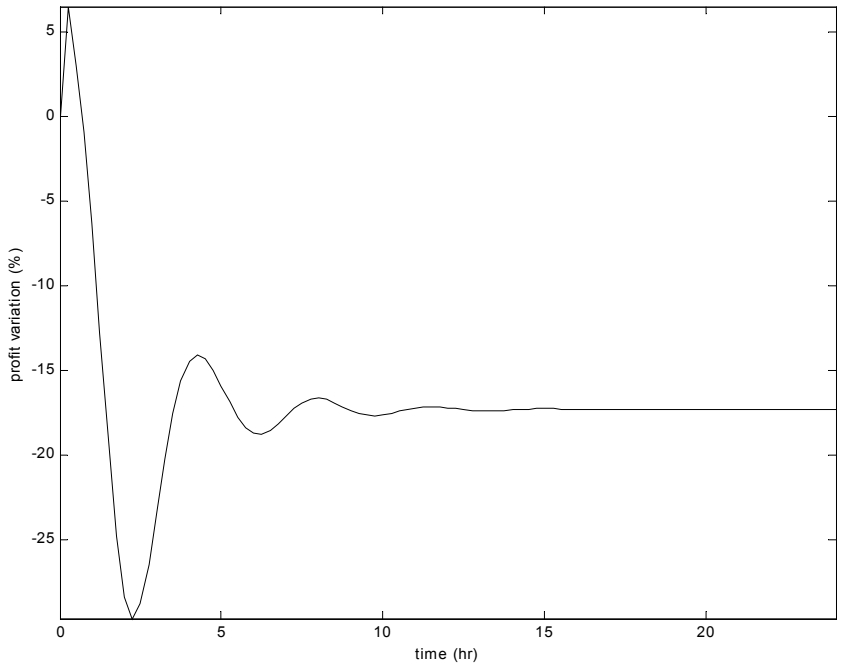


Figure 7.11 Worst-Case Analysis: Profit Profiles in the Second Inner Level

7.7.2 The General Disturbance Characterisation

The second controllability assessment considers the general disturbance profiles, which are taken from historical plant data. The sequential framework developed in Chapter 6 is utilised in this case. The considered disturbance profiles are Q_F, T_F and T_{CW}, while x_F are assumed constant at the nominal value. The disturbance profiles are sampled and held every 30 minutes, and their profiles relative to the plant average are shown in Figures 7.12-7.14.

There are two conditions evaluated within a 24 hour time horizon:
Case 1: Time sequence: 4 hours, time window: 4 hours
Case 2: Time sequence: 24 hours, time window: 24 hours

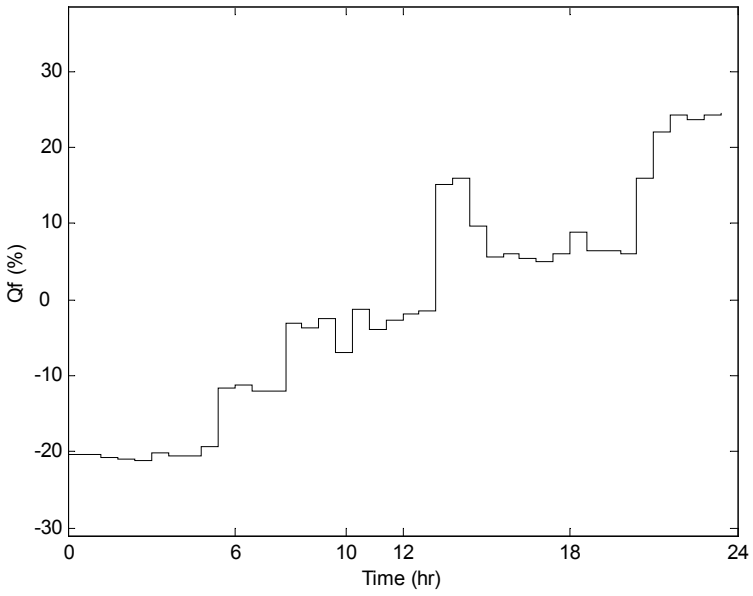


Figure 7.12 Disturbance Profiles: Feed Flowrate Q_F

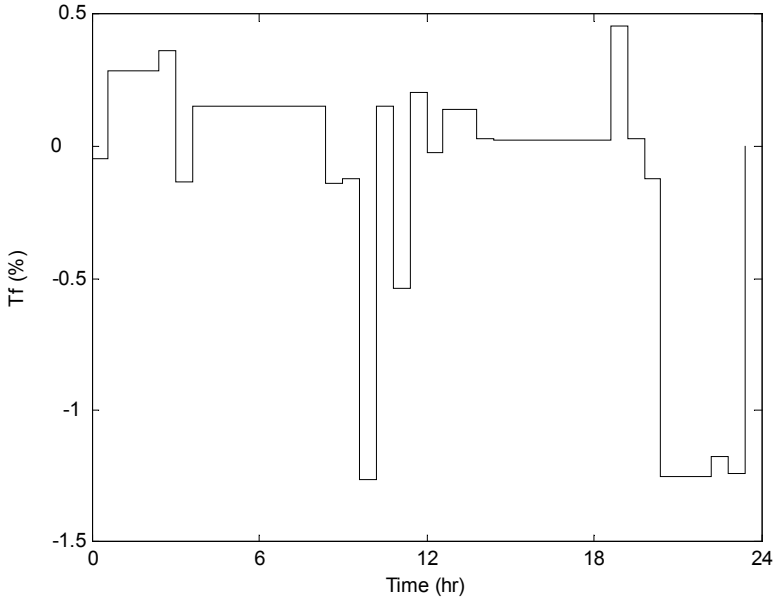


Figure 7.13 Disturbance Profiles: Feed Temperature T_F

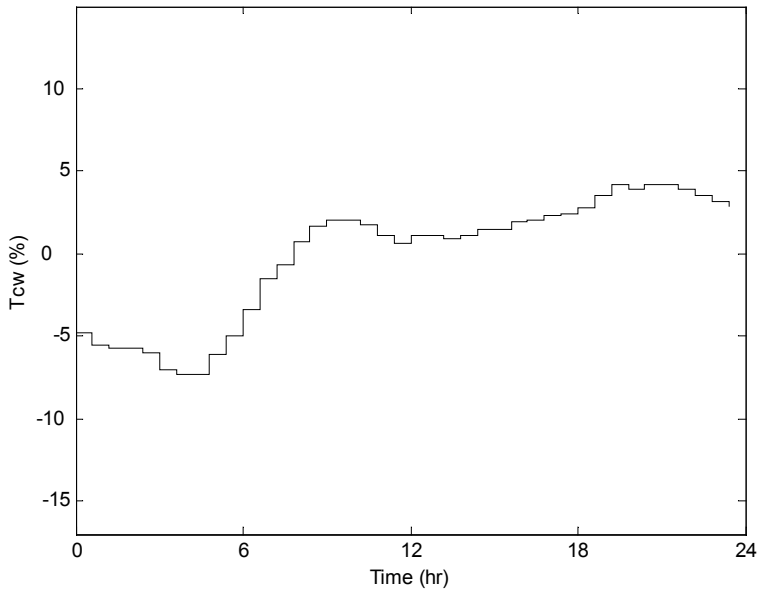


Figure 7.14 Disturbance Profiles: Cooling Water Temperature T_{CW}

These different time windows are used to evaluate the computational costs of solving the piecewise dynamics within the framework. The 24h window is intended for the non-sequential optimisation over whole time horizon and the benchmark solution for the other windows. However, due to the size of the problem, the assessment of this window takes 1392 CPU minutes (23.2 hours) to complete. The results are given in Table 7.8.

Table 7.8 Sequential Controllability Assessment Results

	Steady State Optimum	Closed-loop optimum	
		Case 1	Case 2
Φ	9.12	9.06	9.05
ρ_{P4}	$1.03\rho_{P4}$*	$1.03\rho_{P4}$*	$1.03\rho_{P4}$*
ρ_{P5}	$1.03\rho_{P5}$*	$1.02\rho_{P5}$*	$1.02\rho_{P5}$*
T_{P5}	T_{P5}*	$1.03T_{P5}$*	$1.03T_{P5}$*
Q_{I1}	$1.41Q_{I1}^{min}$	$1.41Q_{I1}^{min}$	$1.41Q_{I1}^{min}$
Q_{I2}	$1.38Q_{I2}^{min}$	$1.38Q_{I2}^{min}$	$1.38Q_{I2}^{min}$
Q_{I3}	0.000	0.000	0.000
Q_{I4}	$1.30Q_{I4}^{min}$	$1.30Q_{I4}^{min}$	$1.30Q_{I4}^{min}$
y_{P3} (m_{S3})	0 (0.000)	0 (0.000)	0 (0.000)
y_{P4} (m_{S4})	1 (93.4% steam)	1 (92.4% steam)	1 (92.0% steam)
y_{P5} (m_{S5})	1 (6.5% steam)	1 (7.2% steam)	1 (7.9% steam)
y_{c1} (α_1)	1 (1.000)	1 (1.000)	1 (1.000)
y_{c2} (α_2)	1 (1.000)	1 (1.000)	1 (1.000)
y_{c3} (α_3)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c4} (α_4)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c5} (α_5)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c6} (α_6)	1 (1.000)	0 (0.000)	0 (0.000)
y_{c7} (α_7)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c8} (α_8)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c9} (α_9)	0 (0.000)	0 (0.000)	0 (0.000)
y_{c10} (α_{10})	1 (1.000)	1 (1.000)	1 (1.000)
y_{c11} (α_{11})	0 (0.000)	0 (0.000)	0 (0.000)
y_{c12} (α_{12})	1 (1.000)	1 (0.000)	1 (0.000)
Variable groups	[h_{P1-5}, ρ_{P4}, T_{P5}] [Q_{P1-4}, M_{S34}, M_{CW}, M_{V5}] [Q_{P5}, ρ_{P5}]	[h_{P1-5},T_{P5}] [Q_{P1-4}, ρ_{P4}, Q_{P5}, ρ_{P5}, M_{V5}]	[h_{P1-5},T_{P5}] [Q_{P1-4}, ρ_{P4}, Q_{P5}, ρ_{P5}, M_{V5}]
$\overline{w}$	h_{P1}, M_{S34}, Q_{P5}	T_{P5}, Q_{P5}	T_{P5}, Q_{P5}
r-OCI	0.00035	1	1
CPU time (m)	2.32	592.611	1392.290

There is a substantial computational difference between the two cases, which is due to the number of disturbance segments within the optimisation windows. The most computational cost of the 4h window is spent in the first sequence. Once the optimum control strategy is found, which is similar to the worst-case results, it is applied to other sequences. Therefore, the second to sixth assessment sequences of case 1 run through the inner levels. Since both cases yield similar process and controller structures, only the dynamic profile of case 2 is presented in Figures 7.15-7.22.

The sequential framework finds the similar process and controller structures with the worst-case analysis. The redundancy analysis, however, is having difficulty in finding the collinear group using 5% tolerance, due to the to usage of dynamic data. The more relaxed tolerance, 10% finally finds the groups, which are similar to the worst-case analysis. This indicates the requirement to further improvement in redundancy analysis.

Due to the similar process and controller structures, the variations in all variables are in the same range with the worst-case analysis. The profit variation in the second iteration are significantly lower than the first iteration, which show that simple control strategy can provide feasible with low variability dynamic performance for the given process economy. The combined disturbances profiles are persistent for the given time horizon, therefore in this case study the variables have not had any chance to settle at a certain steady state values or to a certain recognisable profile. This suggests for the use of longer time horizon for the sequential framework.

Nevertheless, these cases demonstrate the use of the sequential framework using the operational data, and the resulting significant saving in computational cost. It would be interesting to see the trade-off between the saving and the process feasibility, especially for longer time horizon. One interesting case would be the analysis over twelve months period, where the cooling water temperature would change significantly between summer and winter, which may require significant process and controller structure changes. On the other hand, the selection of disturbance sampling time, sequence length and optimisation windows would become more crucial and requires thorough analysis. Such study is recommended for future work.

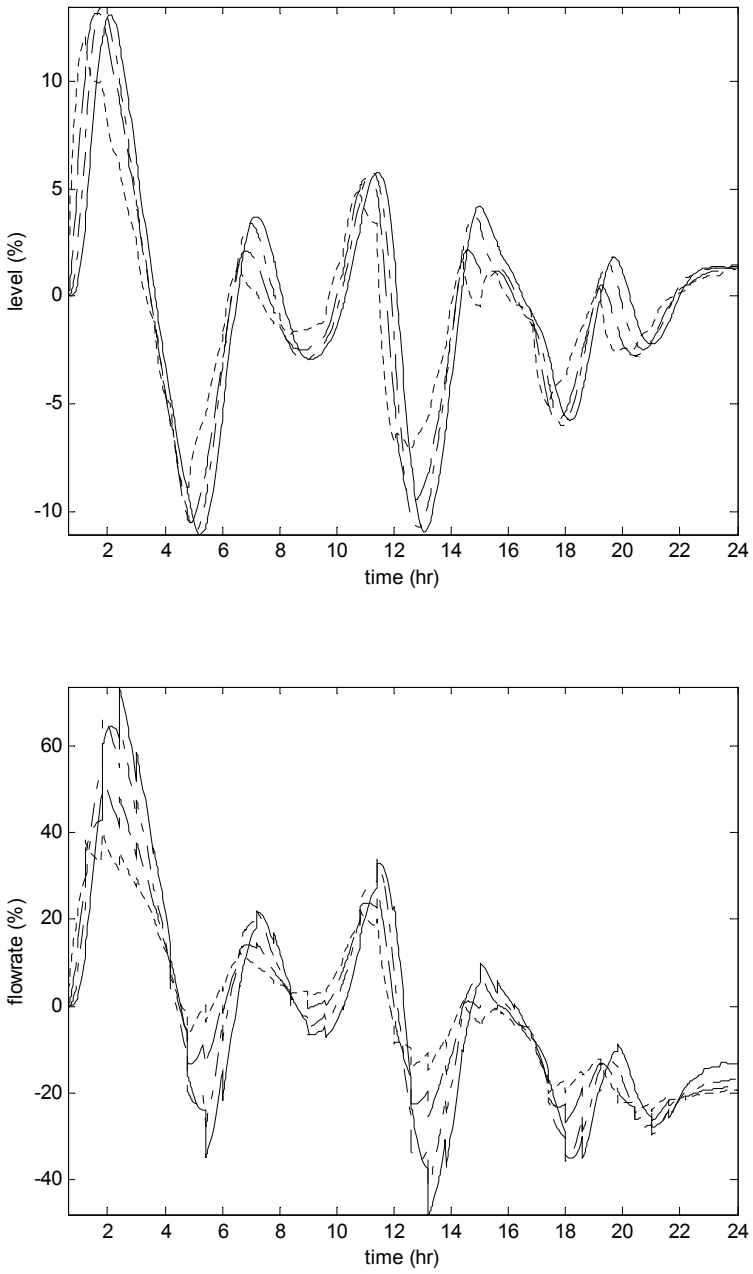


Figure 7.15 General Disturbance Analysis, Case 2:

Level and Product Flowrate Profiles in the First Inner-Level

---- = Stage 1, — — = Stage 2, — - — = Stage 4, — = Stage 5

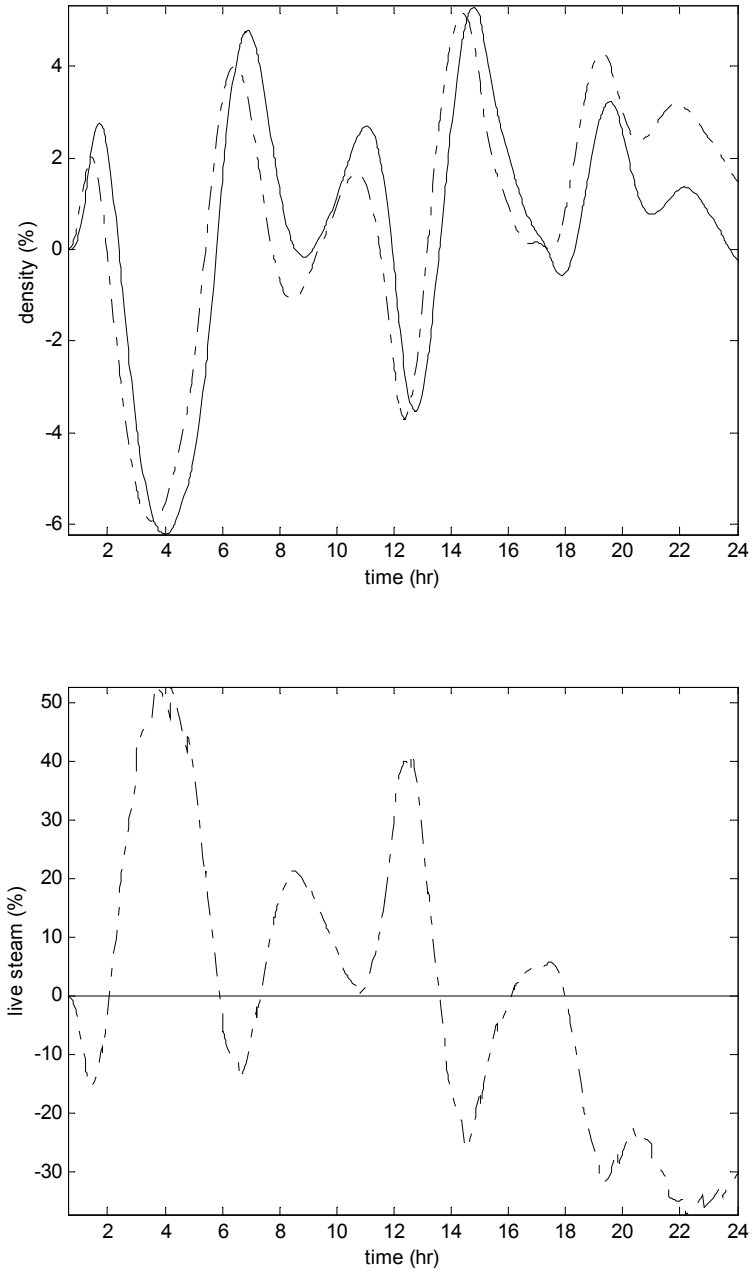


Figure 7.16 General Disturbance Analysis, Case 2:
Density and Live Steam Profiles in the First Inner-Level
— - — = Stage 4, — = Stage 5

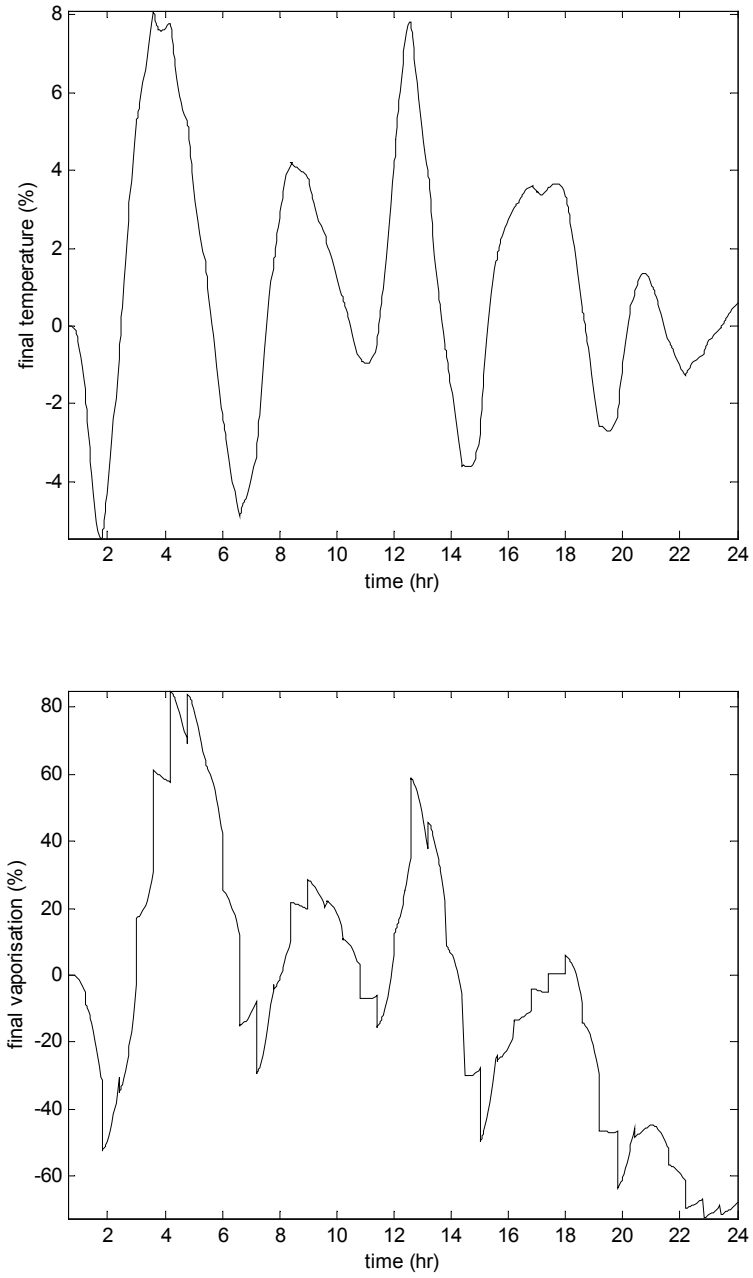


Figure 7.17 General Disturbance Analysis, Case 2:
Temperature and Vaporisation Profiles in the First Inner-Level
— - — = Stage 4, — = Stage 5

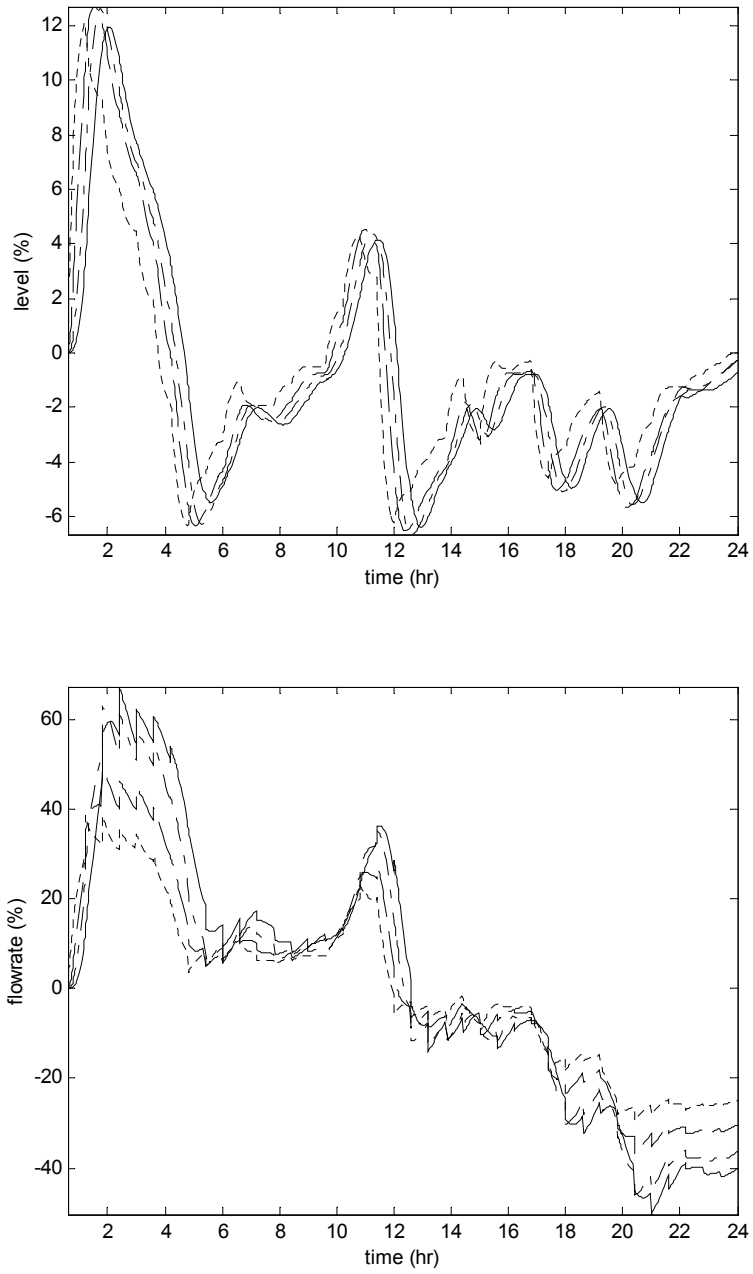


Figure 7.18 General Disturbance Analysis, Case 2:

Level and Product Flowrate Profiles in the Second Inner-Level

---- = Stage 1, — — = Stage 2, — - — = Stage 4, — = Stage 5

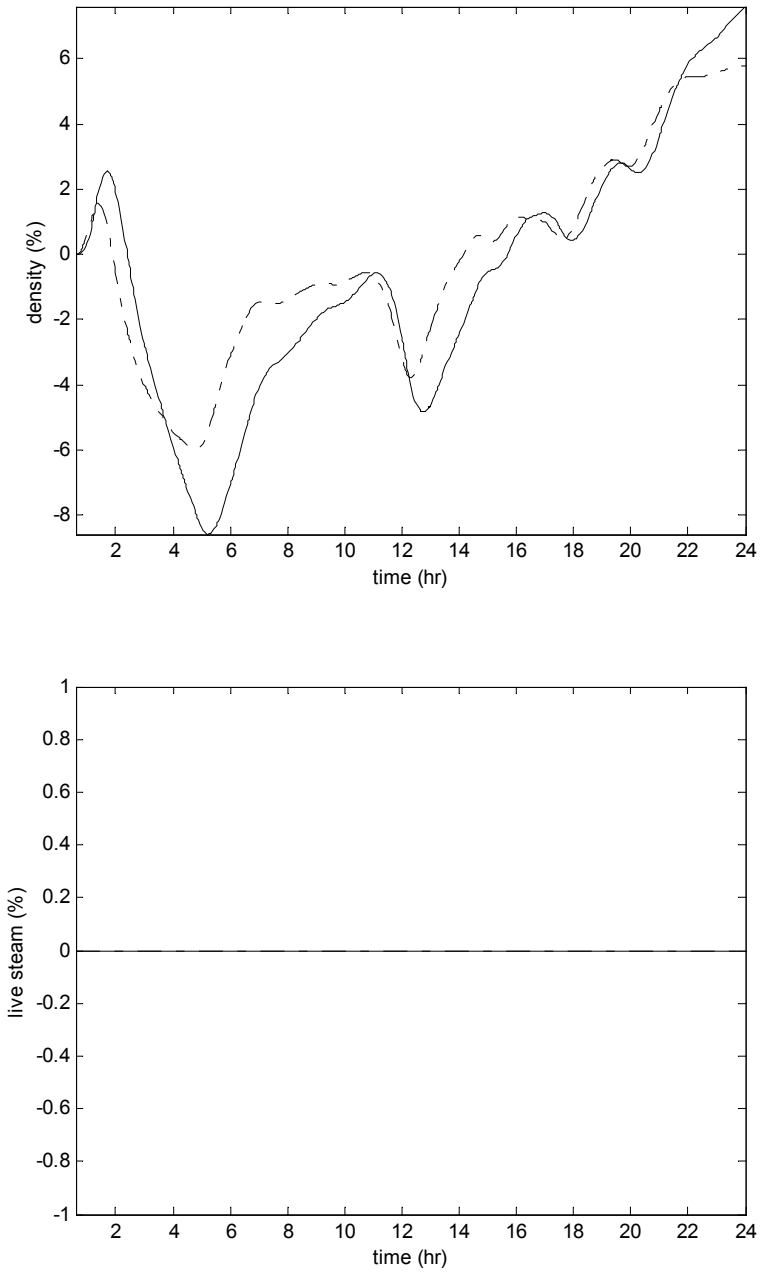


Figure 7.19 General Disturbance Analysis, Case 2:
Density and Live Steam Profiles in the Second Inner-Level
—·— = Stage 4, — = Stage 5

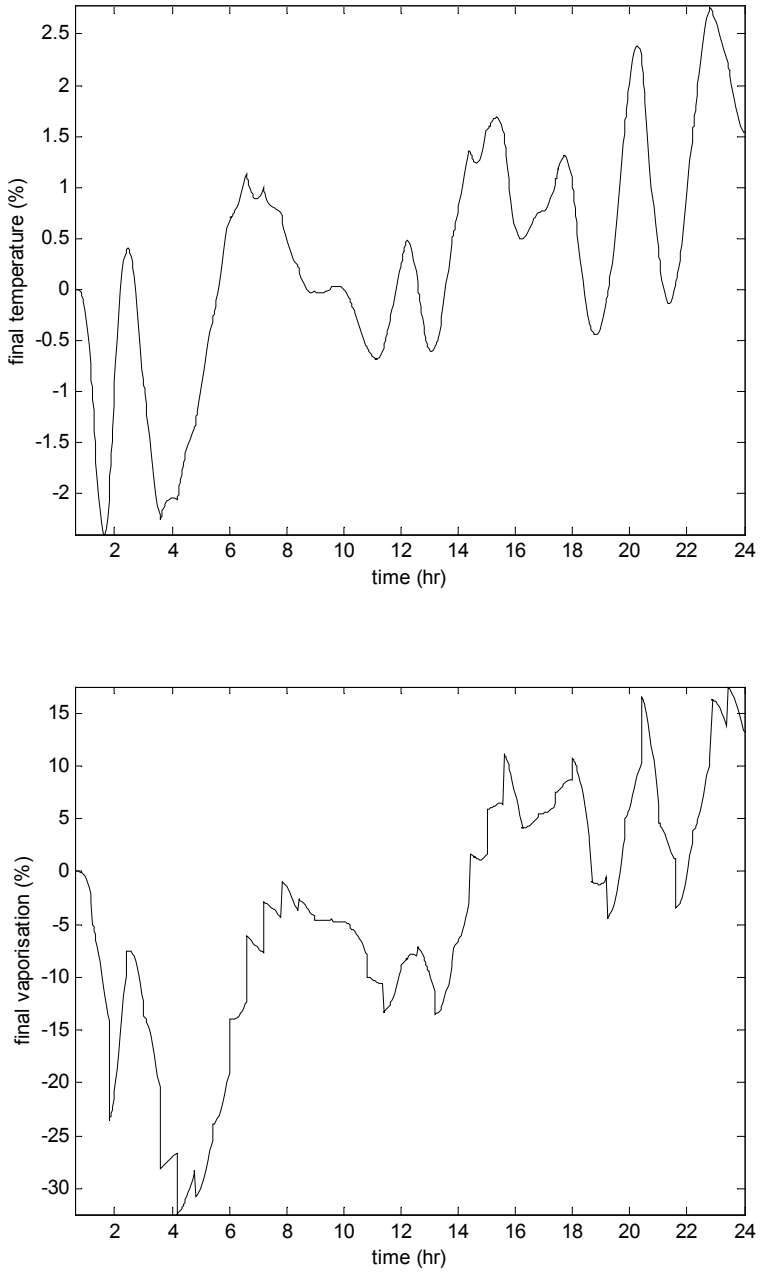


Figure 7.20 General Disturbance Analysis, Case 2:
Temperature and Vaporisation Profiles in the Second Inner-Level
— = Stage 5

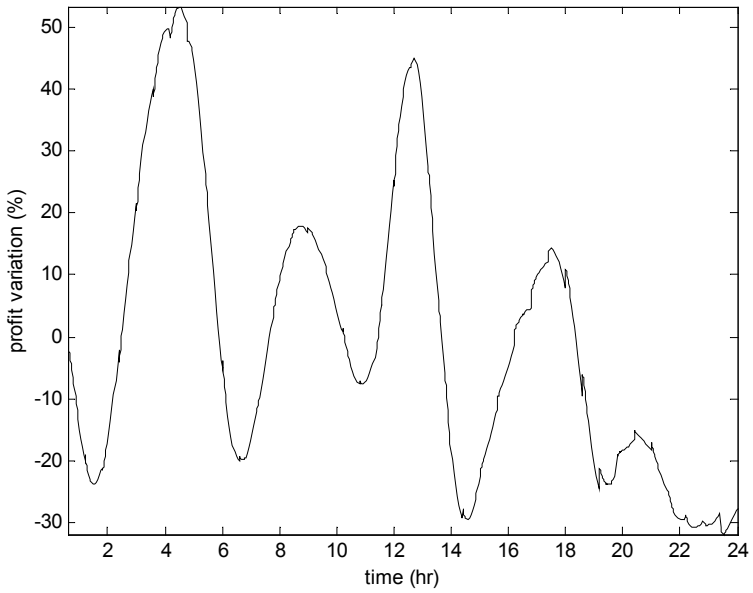


Figure 7.21 General Disturbance Analysis, Case 2:
Profit Profiles in the First Inner-Level

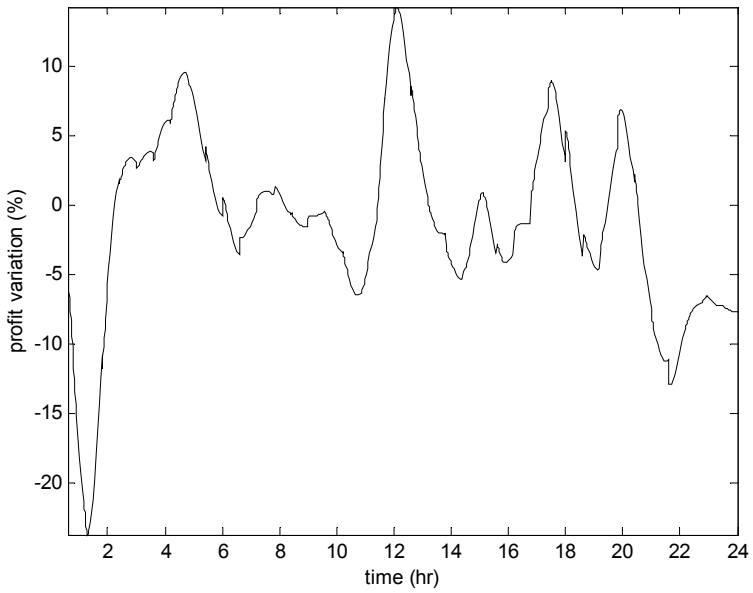


Figure 7.22 General Disturbance Analysis, Case 2:
Profit Profiles in the Second Inner-Level

While the index-2 DAE is solved systematically within the framework, the computational cost is evidently high. This indicates a strong need to develop the efficient solver of high index DAE for application real industrial processes. Therefore, further work to solve high index DAE is recommended.

7.8 Conclusion

The Dynamic Operability Framework (DOF) has been successfully applied to assess the controllability of an industrial five-effect evaporator system at both the worst-case and the general disturbance characterisations. The highest process economy is achieved by a four-effect operation, and simple level and temperature controls have been recommended. The high index DAE problem has been addressed within the framework, highlighting the high computational cost due to strong state dependence. However, the sequential framework proposes substantial savings in computation time when small optimisation windows are used. The trade-off between the window selection and the feasibility of the solution, as well as solution of high index DAE are recommended for future work.

7.9 Nomenclature

Acronyms		
CARIMA	:	Controlled Auto Regressive Integrated Moving Average
DAE	:	Differential Algebraic Equations
DOF	:	Dynamic Operability Framework
GLC	:	Generalised Linearising Control
GMC	:	Generic Model Control
GPC	:	Generalised Predictive Control
IOIMC	:	Input / Output Linearised and Internal Model Control
IOLC	:	Input / Output Linearised Control
LMTD	:	Logarithmic Mean Temperature Difference
r-OCI	:	Regulatory Output Controllability Index
MINLP	:	Mixed Integer Nonlinear Programming
MPC	:	Model Predictive Control

Acronyms		
NLP	:	Nonlinear Programming
OCI	:	Output Controllability Index
ODE	:	Ordinary Differential Equations
PI	:	Proportional – Integral
PVAR	:	Profit Variation
QFT	:	Quantitative Feedback Theory
RGA	:	Relative Gain Array
ROM	:	Relative Order Matrix

Variables and Matrices		
A	:	Flash tank cross sectional area (m^2)
BPE	:	Boiling Point Elevation (°C)
C	:	Solid component (gr/l)
cp	:	Heat capacity (KJ/Kg.K)
E	:	Enthalpies (KJ/Kg)
h	:	Liquor levels (m)
M	:	Mass flow-rates (ton/hr)
P	:	Pressure (KPa)
Q	:	Volume flow-rates (m^3/hr)
T	:	Temperatures (°C)
K_C	:	Controller gain
t	:	Time (hr)
UA	:	Heat transfer coefficient over the cross sectional transfer area
$\overline{\mathbf{w}}$	:	Vector of functional measurement variables
x	:	Solid composition ratio
y_P	:	Binary decision variables defining process structure
y_C	:	Binary decision variables defining controller structure

Greek letters		
α	:	Scaling parameters
θ	:	Disturbances and/or uncertainties
Φ	:	Objective function ($/hr)
Φ_V	:	Objective function for model validation
ρ	:	Densities (Kg/m^3)
τ	:	Controller reset time (hr)

Superscripts		
k	:	Critical values
max	:	Maximum values
min	:	Minimum values
*	:	Target values

Subscripts		
i	:	Stage number
C	:	Condensats
CW	:	Cooling water
F	:	Evaporator main feed
HF	:	Heater feed
HD	:	Heater output
I	:	Flashtank output stream
L	:	Stream tag
P	:	Product stream
S	:	Steam
ss	:	Steady state conditions
V	:	Vapor

7.10 References

Byrnes, C. I. and Isidori, A. (2000). "Output Regulation for Nonlinear Systems: An Overview", International Journal of Robust and Nonlinear Control, 10 (5), 323-337.

Cardoso, A. L. and Dourado, A. (1998). "A Reduced Order Model and a Robust Controller's Synthesis of a Black-Liquor Evaporation System", IFAC Conference on System Structure and Control, Nantes, France.

Chenery, S. D. (1997). "Process Controllability Analysis Using Linear and Nonlinear Optimisation", Ph.D. Thesis, Centre for Process System Engineering and Department of Chemical Engineering, Imperial College of Science, Technology and Medicine, London.

Daoutidis, P. and Kravaris, C. (1992). "Structural Evaluation of Control Configurations for Multivariable Nonlinear Processes", Chemical Engineering Science, 47 (5), 1091 - 1107.

DISR (2000). "Energy Efficiency Best Practice in the Australian Alumunium Industry", Energy Efficiency Best Practice, Department of Industry, Science and Resources, Energy and Environment Division, Canberra.

Ekawati, E. (2002). "Systematic Controllability Assessment in Dynamic Operability Framework", The Third Inter University Postgraduate Electrical Engineering Symposium, Murdoch University, Rockingham Campus, Australia.

Elhaq, S. L., Giri, F. and Ubenhauen, H. (1999). "Modelling, Identification and Control of Sugar Evaporation - Theoretical Design and Experimental Evaluation", Control Engineering Practice, 7 (8), 931 - 942.

EPA (1994). "Increase in Production to 3.3. Million Tonnes Per Annum at Wagerup Alumina Refinery, and Associated Bauxite Mining Activities", Environmental Protection Authority, Perth, Western Australia.

Fabian, G., van Beek, D. A. and Rooda, J. E. (2001). "Index Reduction and Discontinuity Handling Using Substitute Equations", Mathematical and Computer Modelling of Dynamical Systems, 7 (2), 173 - 187.

Glemmestad, B. (1997). "Optimal Operation of Integrated Processes - Studies of Heat Recovery Systems", PhD Thesis, Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway.

Kam, K. M. and Tade, M. O. (1998). "Nonlinear Control of a Simulated Industrial Evaporation System Using a Feedback Linearization Technique with a State Observer", Industrial and Engineering Chemistry Research, 38 (8), 2995 - 3006.

Kam, K. M. and Tade, M. O. (2000). "Simulated Nonlinear Control Studies of Five Effect Evaporator Models", Computers and Chemical Engineering, 23, 1795 - 1810.

Kam, K. M., Tade, M. O. and Le Page, G. P. (1998). "Nonlinear Control of a Simulated Five-Effects Evaporator Using Input-Output Linearization", Industrial and Engineering Chemistry Research, 38 (8), 2995 - 3006.

Kam, K. M., Tade, M. O., Rangaiah, G. P. and Tian, Y. C. (2001). "Strategies for Enhancing Geometric Nonlinear Control of an Industrial Evaporator System", Industrial and Engineering Chemistry Research, 40 (2), 656-667.

Khan, F. I., Gupta, S. C. and Abbasi, S. A. (1998). "Dynamic Modelling and Simulation of Multiple Effect Evaporator System", Hungarian Journal of Industrial Chemistry, 26, 173 - 179.

Kundergi, R. and Nataraj, P. S. V. (1994). "Evaporator Control Design: A Quantitative Feedback Theory Approach", Third IEEE Intl. Conf. on Control Applications, Glasgow, UK.

Liptak, B. G. (1999). "Optimization of Industrial Unit Process", CRC Press LLC.

Mills, P. M., Zomaya, A. Y. and Tade, M. O. (1999). "Application of Adaptive Neural Model-Base Control", 32nd Conference on Decission and Control, San Antonio, Texas, USA.

Montano, A., Silva, G. and Hernandez, V. (1991). "Nonlinear Control of a Double Effect Evaporator", IFAC Advanced Control of Chemical Processes, France, IFAC.

Newell, R. B. and Lee, P. L. (1989). "Applied Process Control: A Case Study", New York, Prentice Hall.

Nielsen, K. M., Pedersen, T. S. and Nielsen, J. F. D. (1996). "Simulation and Control of a Multiple Effect Evaporator", UKACC International Conference on CONTROL'96.

Nooraii, A. (1996). "Kwinana Refinery Mass and Energy Balance", Alcoa of Australia Limited.

Sidrak, Y. L. (2001). "Dynamic Simulation and Control of the Bayer Process. A Review", Industrial and Engineering Chemistry Research, 40 (4), 1146-1156.

Stefanov, Z. and Hoo, K. A. (2003). "A Distributed-Parameter Model of Black Liquor Falling Film Evaporators. Part I. Modeling of a Single Plate", Industrial and Engineering Chemistry Research, 42, 1925 - 1937.

Tjoa, I. B. and Biegler, L. T. (1991). "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential - Algebraic Equation Systems", Industrial and Engineering Chemistry Research, 30, 376 - 385.

To, L. C., Tade, M. O. and Kraetzl, M. (1998a). "An Uncertainty Vector Adjustment for Process Modelling Error Compensation", Journal of Process Control, 8 (4), 265-277.

To, L. C., Tade, M. O., Kraetzl, M. and Le Page, G. P. (1995). "Nonlinear Control of a Simulated Industrial Evaporation Process", Journal of Process Control, 5 (3), 173 - 182.

To, L. C., Tade, M. O. and Le Page, G. P. (1998b). "Implementation of a Differential Geometric Nonlinear Controller on an Industrial Evaporator System", Control Engineering Practice, 6 (11), 1309-1319.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Wang, F. Y. and Cameron, I. T. (1994). "Control Studies on a Model Evaporation Process - Constrained State Driving with Conventional and Higher Relative Degree Systems", Journal of Process Control, 4 (2), 59 - 75.

Winchester, J. and Marsh, C. (1999). "Dynamics and Control of Falling Film Evaporators with Mechanical Vapour Recompression", American Control Conference, San Diego, California.

Young, B. R. and Allen, R. M. (1994). "Gain-Scheduled Lumped-Parameter Multi-Input, Multi-Output Models of a Pilot-Plant Climbing Film Evaporator", Control Engineering Practice, 2 (2), 219 - 225.

Young, B. R. and Allen, R. M. (1995a). "Data-Collection and Identification Strategy for a Climbing Film Evaporator", Measurement & Control, 28 (6), 167-172.

Young, B. R. and Allen, R. M. (1995b). "Multi-Input, Multi-Output Identification of a Pilot-Plant Climbing Film Evaporator", Control Engineering Practice, 3 (8), 1067-1073.

Chapter 8

Conclusion and Future Works

8.1 Introduction

The primary objective of this research has been to develop a systematic controllability assessment tool for process control design, which provides clearer relationships between process profitability, controllability, and operational switching strategies in response to variations in the operating conditions. To achieve this goal, the features of the original Dynamic Operability Framework (Bahri, 1996) has been enhanced through the following stages:

1. Intensive review of the existing methodologies for operability, flexibility and controllability assessments of chemical processes.
2. Integration of a formal controllability index within the Dynamic Operability Framework, and its application to a fixed process control structure.
3. Analysis and elimination of redundant measurement variables to support the computation of controllability index.
4. Development of a dynamic MINLP solver for the efficient selection of optimum process and controller structure.
5. Development of a sequential optimisation scheme to accommodate the controllability assessment subject to the general disturbance profiles.

6. Application of the proposed framework for controllability assessments of an industrial case, which is the five-effects liquor burning evaporator in an Alumina refinery.

These stages and the corresponding achievements are presented in the following section.

8.2 Problems and Main Conclusions

8.2.1 Review of Operability Analyses

An overview of the existing strategies to ensure the operability has been presented, especially on controllability, process control design and process synthesis. It has been shown that controllability is an important consideration in achieving process designs that are economically viable, through the effective use of the available resources in maintaining the design specifications despite the variations in the process operation. This important concern of engineering design is accommodated systematically at the design stage with the use of optimisation algorithms. This include the development of numerical tools to support the assessment of process controllability in terms of parameters and structure, simultaneously with process economy, subject to the affecting disturbances and uncertainties.

The proposed algorithm in this thesis is the extension of the original Dynamic Operability Framework (Bahri, 1996). The original framework is a two-level, iterative optimisation algorithm. This framework addresses the flexibility and controllability of process design, in response to the affecting disturbances and uncertainties, as a semi-infinite, dynamic Mixed Integer Nonlinear Programming (MINLP) problem.

The proposed extensions to this framework are the incorporation of a controllability index that captures the multivariable process dynamics, the efficient selection of optimum process and controller structure, and the accommodation of general disturbance profiles. The modified framework is developed in technical software

MATLAB ver. 6.0 (Mathworks, 2000), and the codes are organised in a toolbox for controllability assessments. These aims has been pursued systematically, as presented in the following five subsections.

8.2.2 Integration of the Output Controllability Index

The Output Controllability Index (Vinson and Georgakis, 2000) has been adapted and integrated within the original Dynamic Operability Framework (Bahri, 1996). The main features of this integration is the geometric representations of the feasible operating range, input and output spaces, which are quantified in terms of their respective convex-hulls. The geometric operations among them show that this flexibility is the special case of the controllability problem. The approach replaces the multiple maximisation problems in the inner-level of the original framework with a sequence of geometric operations, and some of the inequalities in the outer-level with an equality constraint.

The original outcome, which is the best feasible operating condition, has been maintained. The main outcomes of the modified framework are the multivariable controllability index, taking into account all of the measured output variables and their interactions; and the dynamic economic index, representing the variation of process economy due to regulatory dynamics. These are the General Integral Absolute Error (GIAE), and Profit Variation (PVAR), respectively. The application of the modified framework has been demonstrated on the controllability assessment of an exothermic, nonlinear chemical process, at three different but fixed process and controller structures.

8.2.3 Redundancy Analysis and Elimination

The geometric operations of the controllability index, however, may involve collinear output variables. This collinearity creates the redundant dimensions in the output space, which disrupts the geometric computation, and causes incorrect detection of critical disturbance and uncertainty combinations and long computation time.

This problem calls for a mechanism to analyse and eliminate the redundancies of a given set of measured output variables. For this purpose, a redundancy analysis and elimination procedure has been developed based on the Principal Component Analysis (PCA) of the process steady-state correlation data. This algorithm groups the redundant output variables together, and selects one representative functional variable from each group.

This procedure yields the successful geometric operations and correct critical disturbance detection. The capability of finding different functional measured variables, or equivalently eliminating redundant variables, has been shown for the fixed structure, with significant improvement in computation time.

8.2.4 Process and Controller Synthesis

The modified framework has been further extended to cover an automatic and systematic process and controller synthesis. The outer level at the two earlier stages has been considered as a dynamic Nonlinear Programming (NLP) problem. This time, it is tackled as a dynamic MINLP problem. The algorithmic solution of the problem has been developed based on the Branch and Bound method. It is tailored to accommodate the redundancy analysis and elimination procedure to automatically handle the different redundant output variables generated by each process and controller structure candidates. The algorithm performance is enhanced further with a compact MINLP superstructure formulation, the alternating depth-first and breadth-first search, and the rapid filtering of inferior solutions.

The algorithm performance has been demonstrated in four cases of chemical process and controller synthesis. These cases demonstrate fast and efficient convergences at the convex problems, and the best solution of the nonconvex problems. The strong interaction between process and controller characteristics has also been shown, emphasising that controllability depends on both the process and the implemented controller.

8.2.5 Consideration of General Disturbance Profiles

At this stage, the framework has been modified to consider the general disturbance profiles in process controllability assessments. To accommodate this feature, the disturbance profiles are sampled, the corresponding piecewise process dynamics are assessed, and the framework progresses sequentially within an optimisation window over the time horizon. The unfalsification approach has been adopted. Here, the recommended designs for the subsequent sequence is based on the analysis of the sampled disturbances and the associated piecewise process responses within the preceding window. This modification provides more adaptability and reduces conservatism of process design over longer time horizon, and wider disturbance ranges and dynamics. It also allows the assessment to focus on particular disturbance profiles affecting the process, such as those collected from operational history.

The framework performance has been demonstrated in three case studies involving a dynamic nonlinear process in a superstructure. This process is affected by various disturbance signals. The variations of process profitability, controllability, and computational costs due to different optimisation windows have been presented.

8.2.6 Controllability Analysis of an Industrial Five-Effect Evaporator Process

Finally, the framework is applied for systematic controllability assessment of an industrial case. The unit operation under study is the five-effect liquor-burning evaporator associated with the Bayer process of Alumina refinery. The challenge in this process is to handle the strong interaction between level inventories, densities and temperatures. These dynamics are represented by an index-2 Differential Algebraic Equations (DAE) model. The uncertain parameters in this model are estimated based on validation against plant data. The process control design is considered within a superstructure, involving several alternatives of steam supplies and vapor circulation within the process, as well as several alternative pairs of feedback and feedforward control loops. The controllability assessment has been performed successfully at both

the worst-case and the general disturbances profiles, over a 24 hours operational period. Both cases justify the four–effects operational stage, and recommend a simple controller strategy to achieve both the optimum value and the low variability of the process economy.

8.2.7 Overall Achievements

Overall, the Dynamic Operability Framework has been successfully extended to cover the controllability assessment in process control design. The framework works well with general nonlinear process models, capable of selecting the most profitable and controllable process and controller structure, as well as their parameters. The framework generates the indices that provide clearer relationships between process profitability and controllability, and takes account of inherent variable interaction within a multivariable process. It is equipped with an efficient dynamic MINLP solver, dedicated for selecting optimum process and controller structures. It is also capable of considering general disturbance profiles, by switching between the available control strategies in response to variations in the operating conditions. These features have been demonstrated through various case studies, including a real industrial case.

8.3 Recommendation for Future Works

The controllability assessment is an ongoing development. While this thesis provides substantial achievements in the quantifications of process controllability, the solution of process and controller structure selection, and the assessment subject to general disturbance profiles, there are several emerging issues deserving further investigations.

These issues are as follows:

1. The enhanced framework has been focusing on regulatory performances. Therefore, the next logical stage is to expand the capability to handle the servo performance. For this purpose, the framework may be combined with the switchability concept (Vu *et al.*, 1997; White *et al.*, 1985).

2. The objective function of the academic case study presented repeatedly in this thesis has been kept as close as possible with the one in the work of Bahri (1996). The reason is to enable direct evaluation of the algorithms, the computational results and performances. The changes on the objective function to provide better insight to the process, such as including the product concentration and penalties on equipment size or structural changes, are indeed possible. These are therefore recommended for future works.
3. The analysis on computational complexity of the framework can be further enhanced by varying the problem parameters. Such analysis has been reported recently for MILP cases (Till *et al.*, 2003), and a comparable analysis on dynamic MINLP problems would be a valuable complement to this study.
4. The general disturbance profile has been accommodated in the sequential framework. The utilisation of the prediction techniques may further support this performance, by augmenting the predicted process and disturbance information for controller design purpose (Chen *et al.*, 2000; Kam *et al.*, 2001; Mekarapiruk and Luus, 1994; Quintero and Quiennec, 2002; Sun and Tsao, 2001). The trade-off between the window size, computational cost and the degree of feasibility are inherent in this problem and deserve further investigation.
5. In optimising the process dynamics, the framework utilises the dynamic programming approach, which adjusts the initial values based on the calculated final values. It is recommended to investigate the forward dynamic optimisations approach, because it would yield smooth dynamic responses over the time horizon in the sequential framework. This modification may require reformulation of the optimisation problem, which includes constraining the initial values to the final values of the previous sequences, and shaping the dynamic profiles to achieve the optimum objective, with attention to the effective solution of path constraints problem (Bojkov and Luus, 1994; Vassiliadis *et al.*, 1994; Wang and Shieh, 1997).
6. The framework considers the fixed degree, 100% of process flexibility. The consideration of the optimum degree of flexibility would require an efficient

strategy in solving the multi-objective optimisation problem (Schweiger and Floudas, 1997; Wang and Shieh, 1997), which is recommended for future works.

7. The high interconnection between unit operations in industry typically raises high-index DAE problems in process control analysis. The development of an efficient solution to this problem is highly recommended, since it will support the input-output based controller design methodologies and fit better with the real process control conditions.
8. The sequential framework uses the piecewise output profiles in place of the steady state data originally required by the redundancy analysis and elimination procedure. The corresponding collinearity detection becomes more difficult, and requires trial and error to determine the tolerance in collinearity detection of the functional variables. Therefore, further study is required to solve this problem. One possible alternative is the use of differential geometric software with the capability to handle the dimensional redundancy (degeneration) issue (Baotic and Torrisi, 2002; Fukuda, 1996), which is available recently.

8.4 Nomenclature

Acronyms		
DAE	:	Algebraic Differential Equations
GIAE	:	Generalised Integral Absolute Error
MINLP	:	Mixed Integer Nonlinear Programming
NLP	:	Nonlinear Programming
OCI	:	Output Controllability Index
PCA	:	Principal Component Analysis
PI	:	Proportional Integral
PVAR	:	Profit Variation

8.5 References

Bahri, P. A. (1996). "A New Integrated Approach for Operability Analysis of Chemical Plants", Ph. D. Thesis, Dept. of Chem. Eng., University of Sydney, Sydney, Australia.

Baotic, M. and Torrisi, F. (2002). "CDD-MATLAB Interface".

Bojkov, B. and Luus, R. (1994). "Time - Optimal Control by Iterative Dynamic Programming", Industrial and Engineering Chemistry Research, 33, 1486 - 1492.

Chen, X., Fukuda, T. and Young, K. D. (2000). "A New Nonlinear Robust Disturbance Observer", Systems & Control Letters, 41, 189 - 199.

Fukuda, K. (1996). "CDD", Zurich, Switzerland, ETH-Zentrum.

Kam, K. M., Tade, M. O., Rangaiah, G. P. and Tian, Y. C. (2001). "Strategies for Enhancing Geometric Nonlinear Control of an Industrial Evaporator System", Industrial and Engineering Chemistry Research, 40(2), 656-667.

Mathworks (2000). "Using MATLAB Ver. 6.0", Natick, MA, USA, The Mathworks Inc.

Mekarapiruk, W. and Luus, R. (1994). "Optimal Control of Inequality State Constrained System", Industrial and Engineering Chemistry Research, 36, 1686 - 1694.

Quintero, C. S. G. and Quiennec, I. (2002). "State and Disturbance Estimation for an Alternating Activated Sludge Process", IFAC 15th Triennial World Congress, Barcelona, Spain.

Schweiger, C. A. and Floudas, C. A. (1997). "Interaction of Design and Control: Optimization with Dynamic Models", Optimal Control: Theory, Algorithms and Applications, Hager, W. W. and Pardalos, P. M., Kluwer Academic Publishers B. V.

Sun, Z. and Tsao, T. C. (2001). "Control of Linear Systems with Nonlinear Disturbance Dynamics", ACC, Arlington.

Till, J., Engell, S., Panek, S. and Stursberg, O. (2003). "Applied Hybrid System Optimisation: An Empirical Investigation of Complexity", IFAC Conference of Analysis and Design of Hybrid Systems, St. Malo, France.

Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). "Solution of a Class of Multistage Dynamic Optimization Problems, 1. Problems without Path Constraints", Industrial and Engineering Chemistry Research, 33, 2111 - 2122.

Vinson, D. R. and Georgakis, C. (2000). "A New Measure of Process Output Controllability", Journal of Process Control, 10(2), 185 - 194.

Vu, T. T. L., Bahri, P. A. and Romagnoli, J. (1997). "Operability Considerations in Chemical Processes: A Switchability Analysis", Computers and Chemical Engineering, 21(4), S143 - S148.

Wang, F. S. and Shieh, T. L. (1997). "Extension of Iterative Dynamic Programming to Multiobjective Optimal Control Problems", Industrial and Engineering Chemistry Research, 36, 2279 - 2286.

White, V., Perkins, J. D. and Espie, D. M. (1985). "Switchability Analysis", Computers and Chemical Engineering, 20(4), 469 - 474.